ESSAYS ON THE NATURE AND ROLES OF KNOWLEDGE

Nicholas Hughes

A Thesis Submitted for the Degree of PhD
at the
University of Oslo
&
University of St Andrews

2015

Essays on the Nature and Roles of Knowledge

Nick Hughes

A Ph.D thesis to be submitted to the University of Oslo for a double badge degree from the University of Oslo and the University of St Andrews

Department of Philosophy, Classics, History of Art and Ideas
University of Oslo

Department of Philosophy
School of Philosophical, Anthropological, and Film Studies
University of St Andrews

December 2014

1. Candidate's declarations:

I, Nicholas Hughes, hereby certify that this thesis, which is approximately 65,000 words in length, has been written by me, and that it is the record of work carried out by me and that it has not been submitted in any previous application for a higher degree.

I was admitted as a research student in September, 2010 and as a candidate for the degree of PhD in September, 2010; the higher study for which this is a record was carried out in the University of St Andrews between 2010 and 2014.

Date 18/12/2014 signature of candidate

2. Supervisor's declaration:

I hereby certify that the candidate has fulfilled the conditions of the Resolution and Regulations appropriate for the degree of PhD in the University of St Andrews and that the candidate is qualified to submit this thesis in application for that degree.

Date signature of supervisor 18/12/2014

3. Permission for electronic publication: *(to be signed by both candidate and supervisor)*

In submitting this thesis to the University of St Andrews we understand that we are giving permission for it to be made available for use in accordance with the regulations of the University Library for the time being in force, subject to any copyright vested in the work not being affected thereby. We also understand that the title and the abstract will be published, and that a copy of the work may be made and supplied to any bona fide library or research worker, that my thesis will be electronically accessible for personal or research use unless exempt by award of an embargo as requested below, and that the library has the right to migrate my thesis into new electronic forms as required to ensure continued access to the thesis. We have obtained any third-party copyright permissions that may be required in order to allow such access and migration, or have requested the appropriate embargo below.

The following is an agreed request by candidate and supervisor regarding the electronic publication of this thesis:

PRINTED COPY
a) No embargo on print copy

ELECTRONIC COPY
a) No embargo on electronic copy

Date 18/12/2014 signature of candidate signature of supervisor

Please note initial embargoes can be requested for a maximum of five years. An embargo on a thesis submitted to the Faculty of Science and Medicine is rarely granted for more than two years in the first instance, without good justification. The Library will not lift an embargo before confirming with the student and supervisor that they do not intend to request a continuation. In the absence of an agreed response from both student and supervisor, the Head of School will be consulted. Please note that the total period of an embargo, including a continuation, is not expected to exceed ten years.
Where part of a thesis is to be embargoed, please specify the part and the reason

Table of Contents

0. Introduction

This dissertation is comprised of five independent essays on the theme of the nature and roles of knowledge. The essays are intended to be free-standing pieces of work and should be read as such. This introduction first surveys the contents of each essay, before placing them in a broader context and discussing some connections between them. It closes with acknowledgements.

The Essays

The first essay, 'An Existential Argument For Pragmatic Encroachment', offers an argument, conditional on fallibilism about knowledge, for the view that there is pragmatic encroachment on knowledge - that is, that whether or not S knows that p depends not only on broadly truth-related factors, but also on facts about S's practical situation such as how much is at stake for S about being right about p. The argument seeks to force a dilemma on those who wish to deny pragmatic encroachment whilst retaining fallibilism; fallibilists who wish to deny pragmatic encroachment will either have to have to impute massive error on our part about what kind of warrant is needed for an error possibility not to stand in the way of a course of action being justified, or they will have to deny some of the central tenants of Bayesian decision theory. Unlike others found in the literature, this argument doesn't appeal to any universal principles linking knowledge and permissible action, but rather relies only on an existential claim. Pragmatic encroachment theorists, it is suggested, should prefer this argument to those appealing to universal principles because it relies on a logically weaker premise, and in doing so nullifies a number of objections to which previous arguments have been vulnerable. §1 surveys the terrain of the debate. §2 presents the new argument. §3 shows how the denial of pragmatic encroachment faces the dilemma, and discusses some possible responses. §3 concludes that insofar as one wishes to defend pragmatic encroachment, the argument given here provides superior grounds on which to do so, and that possible responses to the dilemma incur substantial theoretical costs.

Recent epistemology has largely turned away from the project of attempting a decompositional conceptual analysis of knowledge and towards investigation of the normative and metaphysical roles that knowledge might uniquely play. To this end, on the normative

side it has been argued that knowledge is, variously, the epistemic norm of belief, assertion, and practical reasoning. On the metaphysical side, Timothy Williamson has argued that one's total evidence consists of only those propositions that one knows to be true, and John Hyman has argued that the fact that p can be one's reason for Φ-ing only if one knows that p. The second essay, 'Environmental Luck Gettier Cases And The Metaphysical Roles Of Knowledge', focuses these two claims about the metaphysical roles of knowledge and argues that they have not been sufficiently well motivated. In both cases, it is argued, the positive arguments for the claims fall short because they conflict with the results delivered by environmental luck Gettier cases. §1 introduces the views of interest and suggests that there are reasons to think that they will stand or fall together. §2 discusses Hyman's thesis linking reasons and knowledge, and argues that it conflicts with the verdicts delivered by environmental luck Gettier cases about the conditions under which S can Φ for the reason that p. §3 discusses Williamson's thesis linking evidence and knowledge, and argues that his two arguments in favour of the E→K direction of the link are undermined by environmental luck Gettier cases. §4 concludes that more work will need to be done if we are to be persuaded by Williamson and Hyman's claims.

The third essay is entitled 'Might The Simulation Heuristic Influence Knowledge Attributions?'. People's knowledge attributions display some intriguing patterns. One of these is that people tend to be less willing to attribute knowledge-that-p to a subject when specific unrealized metaphysical possibilities of error are mentioned, and so made salient, than they are when such possibilities are not mentioned and not made salient, but are present nevertheless. This phenomenon, which I call the 'error-salience phenomenon' has been appealed to to motivate contextualist theses about the semantics of 'knows', and appear to pose a challenge traditional invariantist approaches to the word's semantics. What psychological processes, if any, might explain the error-salience phenomenon? The question has recently generated considerable interest, with Timothy Williamson and John Hawthorne speculatively suggesting that it might be the result of the influence of the availability heuristic, Jennifer Nagel making the case that it is the result of epistemic egocentricity, and Mikkel Gerken arguing that it is the result of epistemic focal bias. Each of these authors has appealed to their proposal to attempt to undermine the force of the apparent challenge that the error-salience phenomenon poses to invariantism about 'knows'. This essay makes a preliminary case that the influence of the *simulation heuristic* may partially explain the phenomenon. §1 introduces the topic. §2 sets the scene by introducing the evidence - both

intuitive and empirical - for the existence of the phenomenon. §3 introduces the simulation heuristic and some of its interesting properties, and makes the case that the phenomenon can plausibly be explained by the hypothesis that knowledge attributions are influenced by the heuristic. §4 compares and contrasts this hypothesis with those offered by Williamson, Hawthorne, Nagel, and Gerken. §5 discusses the implications of the proposed psychological explanations on the debate about the semantics of 'knows'. Some pessimistic remarks are made about the prospects of defending invariantism with such psychological explanations. §6 concludes.

The forth essay, entitled 'Excuses and Epistemic Norms', steps into a debate in current epistemology about the epistemic conditions under which it is permissible to employ p as a premise in one's practical reasoning (or, equivalently, permissible to treat p as certain in one's decision making). It has been argued that it is epistemically permissible to employ p as a premise in one's practical reasoning only if one knows that p (the 'knowledge norm of practical reasoning'). This proposal has faced fierce criticism, with it often being suggested that there are counterexamples to the claim in the form of cases where subjects act on the misleading appearance of conformity with the norm. Those who defend the knowledge norm of practical reasoning attempt to shore up their theory against these would-be counterexamples by appealing to the distinction between being justified and being excused. Subjects who act on the misleading appearance of conformity to the knowledge norm may not be blameworthy or criticisable, it is argued, but this does not show that the relevant norm does not hold, for such subjects are merely to be excused for their transgressions, and excusable transgressions are not counterexamples to the norm. This 'excuse maneuverer' has also met with considerable criticism, and those who reject it take the would-be counterexamples to motivate weaker epistemic norms of practical reasoning demanding, for example, only that subjects reasonably believe that which they act on. Yet it has been suggested that these weaker norms will face would-be counterexamples of precisely the same kind, and thus also require an appeal to the justification/excuse distinction. Do they? This essay explores this question and answers it in the affirmative. The essay is divided into three parts. In §1 the knowledge norm is introduced, the alleged counterexamples it faces are surveyed, and the appeal to the justification/excuse distinction is discussed. It is argued that extant objections to appealing to the distinction to defend the knowledge norm of either unpersuasive or inconclusive. In §2 the weaker proposed norms are surveyed, and some important common features are drawn out. Following this, two ways of arguing for parallel would-be counterexamples to these weaker

norms are presented. Considerations from cases where subjects have misleading higher-order evidence, and considerations from Williamsonian anti-luminosity arguments, are explored as potential generators of would-be counterexamples. It is argued that those who endorse the weaker norms will indeed also have to appeal to the justification/excuse distinction, and that this fact undermines the force of their objections to the knowledge norm. However, it is also argued that we have nevertheless as yet been given no motivation for preferring to draw the distinction at the level of knowledge, rather than somewhere else, and vice versa.

The fifth essay, entitled 'From Moore's Paradox To The Knowledge Norm Of Belief And Beyond', attempts to do two things. Firstly, in §1, to argue that knowledge is the norm of belief in the sense that it is epistemically permissible to believe that p iff one knows that p. And secondly, in §2, to give an account of *why* knowledge, rather than something else, is the norm of belief. The argument of §1 departs from the assumption that it is irrational to believe Moorean conjunctions of the form 'p, but I don't know that p'. From this, it is argued, it follows that it is irrational to believe that p if one knows that one does not know that p. I argue that, given this latter fact, it follows that it is impermissible to believe that p if one does not know that p, because it is only when one has an *excuse* for believing that p without knowing that p - the excuse being that one did not know that one was doing so - that it is rational to do so, and were it permissible to believe that p without knowing that p, one would not need an excuse in order to rationally do so. I conclude that it is permissible to believe that p only if one knows that p. Since it is uncontroversial that it is permissible to believe that p if one know that p, it uncontroversially follows that knowledge is the norm of belief in the sense that it is permissible to believe that p iff one knows that p. With the knowledge norm of belief in hand, §2 addresses the 'why?' question - why it knowledge, rather than something else, the norm of belief? One promising answer to this question, I suggest, can be found from the conjunction of the views that knowledge is fallible safe belief, that belief is a functional kind with constitutive connections with decision making dispositions, and that we employ a heuristic of often treating uncertainties as certainties for the purposes of decision making. Taking uncertainties as certain in one's decision making introduces a kind of epistemic risk to the actions that result from that decision making. Lest the epistemic norms of belief license foreseeably unhelpful decision making procedures, it is argued, we should expect there to be restrictions on when it is permissible to take uncertainties as certain. Conforming to a rule that states that one is permitted to believe that p iff one knows that p (and thus, given the constitutive connections between belief and decision making dispositions, permitted to treat

uncertainties as certain in one's decision making iff one knows that p), is, I argue, a good policy to have, given our needs and limitations. And this explains, at least in part, why it is permissible to believe that p iff one knows that p. §3 considers and replies to objections. §4 concludes.

The Broader Context

The publication in 2000 of Timothy Williamson's 'Knowledge and Its Limits' has proven to be a landmark event in contemporary epistemology. I have heard it remarked that prior to the publication of that book epistemology had acquired a reputation for being a rather stagnant area of philosophical theorizing, obsessed with a degenerative research programme of offering ever more convoluted analyses of knowledge in terms of more basic constitutive parts in an effort to solve 'the Gettier problem'. Whether or not that is a fair characterisation, the novelty of Knowledge and Its Limits was to turn this approach on its head. Knowledge, Williamson argued, is not to be understood in terms of more basic constitutive parts, but instead to be taken as a primitive in our theorising, and other philosophically interesting phenomenon understood in terms of it. This 'knowledge-first' inversion of the traditional order has resulted in a proliferation of work exploring new avenues of epistemological thought, and several of the essays in this dissertation can be seen as efforts to contribute to this ongoing exploration. Most obviously the essays 'Environmental Luck Gettier Cases And The Metaphysical Roles Of Knowledge', 'Excuses And Epistemic Norms', and 'From Moore's Paradox To The Knowledge Norm Of Belief And Beyond'. Each of these essays focuses on questions about what normative or metaphysical roles knowledge might play - on the normative side as providing the conditions for permissible belief and practical reasoning, and on the metaphysical side, as being the condition under which one has p as evidence, and can Φ for the reason that p. As the reader will see, I am rather more optimistic about the normative role of knowledge than I am about the metaphysical roles.

Not all the work in this dissertation can be traced back to discussions inspired, or reinvigorated, by Knowledge and Its Limits. The impetus for answering the question of what psychological mechanisms might explain the error-salience phenomenon, discussed in 'Might

The Simulation Heuristic Influence Knowledge Attribution?', finds its origins in debates about the semantics of 'knows' that have been ongoing since early in the second half on the 20th century. Debates which have, in turn, been largely driven by a desire to answer sceptical concerns about the possibility of knowledge first posed by Descartes. The view defended in 'An Existential Argument For Pragmatic Encroachment'; that knowledge is pragmatically encroached upon, also first came about, at least in part, as a result of this debate (though the early development of the view by Fantl and McGrath doesn't fit neatly into this conception of history).

These five essays are some of many written over the course of the last four years I have spent grappling with epistemological issues. The earliest of them - 'An Existential Argument For Pragmatic Encroachment' - was largely written across late 2010 and early 2011. In contrast, the most recent - 'From Moore's Paradox To The Knowledge Norm Of Belief And Beyond' - was written almost entirely in late 2014. Although the reader may occasionally get the opposite impression, I have very few, if any, settled opinions about the matters that this dissertation discusses. The essays are best viewed as a series of snapshots of an ongoing process of investigation. Of course, one of the disadvantages of such a diachronic series of snapshots is that it is liable to reveal tensions between one's earlier and more recent thought. Here I mention three.

I express some pessimism towards (if not outright rejection of) the equation of evidence and knowledge in 'Environmental Luck Gettier Cases And The Metaphysical Roles Of Knowledge', yet I also argue that knowledge is the norm of belief in 'From Moore's Paradox To The Knowledge Norm Of Belief And Beyond'. But there may be reason to think the former motivates the latter. It is rather odd to think that one could have p as evidence without being epistemically permitted to believe that p, and similarly odd to think that one could be epistemically permitted to believe that p without having also having p as evidence for further inquiry. This is one of many lines of thought that I have been unable to pursue in the dissertation, but hope to in future work. Another tension might be thought to exist between the argument I offer in favour of pragmatic encroachment, and the explanation I give of why knowledge is the norm of belief. The former relies on some tenets of traditional Bayesian decision theory that the latter, one might think, gives us reason to be suspicious of. This too is an issue I hope to investigate in future work. Lastly, in 'Excuses and Epistemic Norms' I suggest that we have thus far been given no reason to draw the distinction between justifiably

employing p in one's practical reasoning and excusably doing so at the level of knowledge, rather than mere reasonable belief. But if it is permissible to believe that p iff you know that p, then given certain plausible conceptions of the nature of belief, it follows that the distinction should in fact be drawn at the level of knowledge. Again, I hope to investigate this issue in future work.

No doubt there are other tensions that I have failed to recognise. Nevertheless, I hope that these essays manage, in a small way, to push forward some lines of epistemological inquiry.

Acknowledgements

The essay 'Environmental Luck Gettier Cases And The Metaphysical Roles of Knowledge' is comprised of two separate articles: 'Consistency and Evidence', and 'Is knowledge the ability to Φ for the reason that p?' which are published in Philosophical Studies and Episteme respectively. I thank Springer and Cambridge University Press for giving me permission to reproduce them here.

I have discussed the ideas in these essays with many people. Thanks to Derek Ball, Sebastian Becker, Einar Bohn, Buddy Boren, Cameron Boult, Mark Bowker, Daniel Brigham, Katharine Browne, Wesley Buckwalter, Jamie Burns, Lawrence Carrick, Yuri Cath, Laura Celani, Stewart Cohen, Timothy Chan, Lars Christie, Michael De, Josh Dever, Paul Dimmock, Daniel Drucker, Andy Egan, Jola Feix, Suki Finn, Gareth Fitzgerald, Christel Fricke, Noah Friedman-Biglin, Froydis Gammelsaeter, Jie Gao, Mikkel Gerken, John Gibbons, Ephraim Glick, Jeremy Goodman, Daniel Greco, Patrick Greenough, Steven Hall, Matthew Hancock, Michael Hannon, Hedda Hassel Morch, John Hawthorne, Stephen Hetherington, Thomas Hodgson, Torfinn Huvenes, Jennifer Hornsby, Bruco Jacinto, Carrie Jenkins Ichikawa, Jonathan Jenkins Ichikawa, Daniel Kahneman, Rosanagh Ker, Dirk Kindermann, James Kirkpatrick, Sebastian Koehler, Katrin Kvaran, Dan Labriola, Julia Langkau, Jon Litland, Clayton Littlejohn, Brent Madison, Anna-Sara Malmgren, Brain McElwee, Aidan McGlynn, Matthew McKeever, Robin McKenna, Tomas Midttun Tobiassen, Jennifer Nagel, Anders Nes, Dilip Ninan, Andrea Onofri, Carlotta Pavese, Andrew Peet, Jessica Pepp, David Plunkett,

Duncan Pritchard, Chris Ranalli, Francois Recanati, Monica Roland, Bernhard Salow, Arthur Schipper, Anders Schoubye, Sascha Settegast, Patrick Sherrif, Nico Silins, Will Small, Martin Smith, Jason Stanley, Rachel Sterken, Andreas Stokke, Margot Strominger, Josh Thorpe, Michael Traynor, Leo Townsend, Nick Treanor, Marissa Wallin, Jennifer Wang, Sebastial Watzl, Brian Weatherson, Ralph Wedgewood, Emilia Westlin, Sam Wilkinson, Timothy Williamson, Jonnie Wolf, and Juhani Yli-Vakkuri, for helpful discussions. Thanks also to anonymous referees from Philosophical Studies and Episteme, to audiences at the University of British Columbia, the University of Cambridge, the University of Edinburgh, the University of Manchester, the University of Oslo, and the University of St Andrews, and to all the administrators at Arche and CSMN. Apologies to anyone I have forgotten.

Special thanks to my supervisors Matthew McGrath and Olav Gjelsvik, and in particular to my co-first-supervisors supervisors Herman Cappelen and Jessica Brown. I am enormously grateful for their ongoing encouragement, advice, guidance, and feedback. I find it hard to imagine that there could be two better mentors than Herman and Jessica.

This work is dedicated to those to whom I owe the greatest debt of gratitude: my family Richard, Jayne, and Jennifer, Hughes.

1. An Existential Argument For Pragmatic Encroachment

I. Pragmatic Encroachment

A traditional view in epistemology has it that whether an individual's true belief that p counts as knowledge or not depends only on truth-directed factors such as evidence, sensitivity to truth across possible worlds, reliability of belief forming process, and coherence consideration such as whether the individual has doxastic defeaters. On this view, considerations from the domain of practical rationality have direct no bearing on the possession, or otherwise, of knowledge. The traditional view has recently been challenged. Jason Stanley (2005), John Hawthorne (2004), and Jeremy Fantl & Matt McGrath (2002, 2007, 2009) have all argued that there is 'pragmatic encroachment' on knowledge. That is, whether an individual knows that p is determined, in part, by practical facts about the individual such as the consequences for them of being wrong about p. Thus far, pragmatic encroachment (hereafter 'PE') has been primarily defended by appealing to 'knowledge-action principles'.[1] Such principles purport to capture important connections between epistemic states and practical rationality in the form of epistemic norms of practical reasoning. A number of different principles have been proposed. What they all have in common is that each claims that knowing that p is either a necessary or sufficient condition (or both) for a subject to be epistemically justified in acting on p, or relying on p in their decision making. Stanley (2005) and Hawthorne (2004) argue that knowledge is necessary, Fantl and McGrath (2009) argue that knowledge is sufficient, and Stanley and Hawthorne (2008) argue that knowledge is necessary and sufficient. Another thing the principles have in common is that they all make universal claims; each principle is claimed to hold for all subjects in all worlds at all times.

These authors have used their knowledge-action principles to argue for PE. Fantl and McGrath (2009) argue that if fallibilism is true - if a subject can know that p even though p has an epistemic probability, for the subject, of less than 1 - then PE must also be true. The

[1] Another way that pragmatic encroachers have argued for their view is to present pairs of cases where the subject has the same evidence in each case, but much more is at stake for her about being right in one case than the other. They then argue that, intuitively, the subject knows that p in the low stakes scenario, but not in the high stakes scenario, and infer PE from this. This sort of argument can be found in Stanley (2005), amongst others. I won't have anything to say about such arguments here, except to note that empirical investigation has not yet returned a settled view on the claims about intuitiveness made in the course of the arguments. See, e.g. Schaffer & Knobe (2012), Stanley & Sripada (2012) et al.

argument goes like this:[2] if fallibilism is true, and it is true that as the practical stakes rise a subject needs a stronger epistemic position with respect to p in order to be justified in acting on p, then there exist case pairs which differ only in the stakes in which a subject knows that p in the lows stakes scenario, but is not in a good enough position with respect to p to be justify in acting on it in the high stakes scenario. Given that knowledge that p is sufficient for a subject to be justified in acting on p, it follows that the high stakes subject does not know that p. Since the low stakes subject does know that p, and the only difference between the high and low stakes subjects is in the practical stakes, it follows that knowledge depends, in part, on practical stakes, and hence that PE is true. Brown (2013) suggests that one can also argue for the same conclusion using a necessity direction knowledge-action principle. In this case the argument will be abductive, claiming that the best explanation of why the subject's warrant for p is not good enough for her to be justified in acting on p is because she fails to meet a necessary condition on being so justified; she does not know that p. This sort of argument may be at work in Hawthorne (2004) and Stanley (2005).

These arguments have generated a great deal of discussion, much of it negative. Jessica Brown (2008), Ram Neta (2009), Baron Reed (2010), Stephen Schiffer (2007) and Mikkel Gerken (2011) all argue that the relevant knowledge-action principles are subject to counterexamples, and hence false. Their falsity, it is claimed, significantly undermines the case for PE. PE theorists have offered responses to the apparent counterexamples, and there is an ongoing debate over their status.[3] What seems to have gone largely unnoticed so far, however,[4] is that one needn't appeal to universal knowledge-action principles to defend PE. PE theory does not claim that practical consequences always affect knowledge, but rather that they can sometimes do so. As long as there are at least some possible cases of practical consequences affecting knowledge in the relevant way, PE is true.[5] In theory, the view can be defended with an existential argument. I'm going to claim that, in this case, what is true in theory is also true in practice.[6]

[2] The reconstruction of Fantl and McGrath's argument that I'm about to present is borrowed from Brown (2008).
[3] This issue is discussed in detail in the forth essay of this dissertation - 'Excuses And Epistemic Norms'
[4] An exception is Weatherson (2011, 2012)
[5] As Weatherson (2011) points out, one could hardly claim that practical consequences always affect knowledge, since when one has no evidence whatsoever for the truth of p one will not know that p, irrespective of the practical circumstances.
[6] The argument of this essay isn't one that can motivate invariantist PE theories over contextualism, or vice versa. But it is compatible with both.

If I'm right, and PE theory can be defended with an existential argument, this has important ramifications for the debate. Much ink has been spilled arguing for and against the proposed knowledge-action principles. However, whilst the question of just what epistemic norms govern practical rationality is an interesting one in its own right, for the purposes of establishing the truth or falsity of PE the debate is unnecessary. Let the proposed knowledge-action principles all be false, or let there be no universal knowledge-action principles; this doesn't imperil the case for PE when it is argued for using an existential. Counterexamples to the universal principles don't affect the existential argument: provided that there is at least one case of stakes affecting knowledge in the relevant way, PE is true. Moreover, the kind of existential I will use to argue for PE is entailed by, but does not entail, the knowledge-action principles Fantl, McGrath, Stanley and Hawthorne defend. In defending PE theory using universal principles, these authors take on additional theoretical commitments that are unnecessary to establish their conclusions. Insofar as these additional commitments are controversial, extant arguments for PE have been needlessly mired in controversy. Accordingly, PE theorists ought to prefer arguing from the existential to arguing from universal principles. In doing so, they rely on a logically weaker, and so dialectically stronger, premise. One that isn't subject to the criticisms found in the literature.

II. An Existential Argument For Pragmatic Encroachment

So, suppose that there are case pairs that differ only in the stakes that have the following three features: (a.) the subject knows that p in the low stakes scenario. (b.) if she knows that p in the high stakes scenario, then p is warranted enough to justify the subject in Φ-ing the high stakes scenario, and (c.) p isn't warranted enough to justify her in Φ-ing the high stakes scenario.[7] If

[7] The locution 'p is warranted enough to justify Φ-ing' is borrowed from Fantl and McGrath (2009). A proposition is warranted enough to justify you Φ-ing, according to Fantl and McGrath, when no weakness in your epistemic position with respect to p s*tands in the way* of p justifying you in Φ-ing. That is: your epistemic position with respect to p is strong enough that p can justify you in Φ-ing. That doesn't necessarily mean that p *will* justify you in Φ-ing. When there isn't an appropriate connection between p and

such case pairs are possible, then PE is true. The low stakes subject knows that p. Given that, if she knows that p in the high stakes scenario, p is warranted enough to justify the high stakes subject in Φ-ing, if p isn't warranted enough to justify her in Φ-ing in the high stakes scenario, it follows that the high stakes subject doesn't know that p. Since the cases differ only in the stakes, and there is a difference in knowledge between the low stakes subject and the high stakes subject, it follows that practical consequences can affect knowledge in the relevant way, and that PE is true. Notice that the previous arguments given for PE in the literature - those that argue from knowledge-action principles - entail that such case pairs must be possible. But the converse does not hold: the possibility of such case pairs does not entail that the relevant knowledge-action principles are true, or that those arguments succeed. In this sense, the existential claim that such case pairs are possible make for a logically weaker, and hence dialectically stronger, premise in an argument for PE.

Could such case pairs exist? I think they could. In the proceeding argument, I'm going to assume that a particular kind of fallibilism is true of knowledge; that a subject, S, can know that p when p has an epistemic probability of less than 1 for S. So the conclusion of the argument is going to be a conditional: if fallibilism is true, then so is PE. Of course, some will reject the antecedent. It's not feasible to get into the fallibilism/infallibilism debate here. But fallibilism is a widely enough held view that the conclusion ought to be of considerable

---

Φ-ing, your warrant for p, even if it is strong enough to justify you, won't justify you, because p isn't the kind of thing that can justify Φ-ing. 'Φ-ing' includes believing, doing, preferring, wanting, liking, hating, and intending, amongst other things. Roughly speaking, anything one can do that can be evaluated in terms of it's rationality. One might wonder just what it means to say a subject is justified or unjustified in Φ-ing. Fantl and McGrath (2009) conceive of things as follows: just as it makes sense to ask whether someone's belief that p is justified or unjustified, so it makes sense to ask whether you are epistemically justified in doing some action, or having some attitude, Φ. If someone believes that p without any evidence then their belief is epistemically criticisable: they lack sufficient warrant to be epistemically justified in believing that p. Similarly, if someone performs some action, or holds some attitude Φ, their motivating reason for Φ-ing being that they believe that p, if they have no evidence for the truth of p, then their action or attitude is epistemically criticisable: they lack sufficient warrant for p for it to justify them Φ-ing. Here's an example. Suppose an onlooker sees me selling my lottery ticket to a friend for 10p. The onlooker may ask why I am selling my ticket. If I cite my belief that my ticket will lose as my motivating reason for selling the ticket, then my belief in the truth of the proposition 'my ticket will lose' becomes open to epistemic evaluation in connection to my action of selling the ticket. Suppose I've got no evidence that my ticket is a losing one. Then one way the onlooker can criticise my action - my selling the ticket - is by pointing out that I've got no good reason to believe that my ticket will lose. If they do criticise me in that way, they are offering an epistemic criticism of my action - they are pointing out that I am not well enough epistemically positioned with respect to the proposition 'my ticket will lose' for it to justify me selling the ticket for 10p. Equivalently, p isn't warranted enough for me to justify me selling the ticket. Conversely, if I somehow know that my ticket will lose (say I know that the lottery is rigged), then I can respond to the onlooker's criticism by saying: 'actually, I know that my ticket will lose', and thereby nullify their criticism. When I am in a sufficiently good epistemic position with respect to a proposition p, p is warranted enough to justify me Φ-ing. And if p is the kind of that can conclusively justify Φ-ing, then p justifies me Φ-ing.

interest nevertheless.[8]

So consider the following: Sarah is told by Rosanagh, who is an honest woman, and a reliable testifier on such matters, that the scheduled date of the next Nico Muhly London concert is April 10th (hereafter let 'p' mean 'the scheduled date of the next Nico Muhly London concert is April 10th'). Sarah doesn't care whether p is true or not, and isn't planning to do anything with the information. Nevertheless, she knows that Rosanagh is honest and a reliable testifier, so she forms the belief that p. It turns out that p is true, and that Rosanagh knows that it is. Given these things - and assuming, as we are, fallibilism about knowledge - it is reasonable to think that Sarah knows that p. We are often happy to ascribe knowledge to people who form their beliefs on the basis of reliable but fallible testifiers, and this is what Sarah has done. Fallibilists should accept that, ceteris paribus, Sarah can know that p. Denying this would have unappealingly sceptical ramifications.

Now suppose that Eleanor has the same evidence as Sarah for the truth of p - Rosanagh also told her that p - but that her practical situation is different. Coincidently, just before she spoke to Rosanagh, Eleanor ran into a friend who is fond of low-stakes gambling. The friend offered Eleanor a gamble on whether p. She hasn't decided whether to accept it yet, and now that she has Rosanagh's testimony in hand, she can still do so. Eleanor's stakes are as follows: if she takes the gamble, and p is true, she wins 10p. But if she takes the gamble and p is false, she loses £5. I think the following two things are true of Eleanor (I'll defend them in the next section by arguing that denying them has unpalatable consequences). Firstly, p is not warranted enough for Eleanor to justify her accepting the gamble. Secondly, although Eleanor's stakes are not great, nor are they extremely unfavourable, so it is plausible that, were Eleanor to know that p, p would be warranted enough to justify her accepting the gamble. If these two things are true it follows that Eleanor doesn't know that p.

So I think the following three things are true of Sarah and Eleanor's case:

(1) Sarah knows that p

[8] Fantl and McGrath (2009) likewise offer this conditional as their conclusion. Stanley (2005), and Hawthorne (2004) both develop their arguments for PE from within a Williamsonian infallibilist framework. I believe that the argument I provide in this essay can be adapted for a Williamsonian epistemology, but it isn't feasible to undertake this project here.

(2) If Eleanor knows that p, p is warranted enough for Eleanor to justify her accepting the gamble

(3) Given her stakes, p isn't warranted enough to justify Eleanor accepting the gamble.

The conjunction of (2) and (3) entails that Eleanor doesn't know that p. By (1), Sarah does know that p. As their respective situations differ only in the practical stakes, if these three claims are true of Sarah and Eleanor, then the case pair instantiates the state-of-affairs I described at the beginning of this section (albeit on an interpersonal level), and PE is true.

III. The Traditionalist's Dilemma

I'm going to take it as a given that (1) is true. We are assuming fallibilism, and on fallibilism it ought to be unproblematic to think that Sarah knows that p - her's is a perfectly normal case of fallible knowledge, if there is such a thing. So traditionalists will have to deny that the conjunction of (2) and (3) is true. They can either deny that knowledge that p would be sufficient for p to be warranted enough to justify Eleanor accepting the gamble. ('option 1'), or deny that p isn't warranted enough to justify Eleanor accepting the gamble ('option 2'). Both options, I will argue, incur substantial costs. Let's look at them in turn.

Option 1: Deny that knowledge is sufficient for Eleanor

The first option for traditionalists is to argue that even if Eleanor knows that p, given her stakes, p still isn't warranted enough to justify her accepting the gamble. If this is right then the case pair doesn't entail that Eleanor doesn't know that p, and hence doesn't force the conclusion that practical stakes can affect knowledge. I don't think this is a promising option

for traditionalists, simply because the claim is implausible. Some commentators (e.g. Brown (2008)) have presented cases where knowledge that p seems to be insufficient for p to be warranted enough to justify a subject in Φ-ing. But these cases involve stakes that are far higher than those Eleanor faces. In one of Brown's cases Liz wins £1 if she is right, but loses her home if she is wrong. In another a surgeon saves a few seconds if she is right, but runs the risk of doing irreparable harm to, or even causing the death of, a patient if she is wrong. However, whilst it would be unfortunate for Eleanor to lose £5, and whilst such a loss is more substantial than a gain of 10p, the loss of £5 couldn't by any stretch of the imagination be called significant (assuming, as we will, that Eleanor isn't very poor). Eleanor's situation does not have the grave character of those of Brown's subjects. Brown's examples might show that knowledge that p is not always sufficient for p to be warranted enough to justify a subject in Φ-ing, but they don't motivate the claim that knowledge is not sufficient for Eleanor.[9] Moreover, we can provide a positive case in defence of the claim that knowledge would be sufficient for Eleanor by considering others cases.

Firstly, there seem to be many cases where the negative consequences of being wrong are much greater than they are for Eleanor, but where knowledge is nevertheless intuitively sufficient. Consider: Sarah is running late for lunch date so she needs to drive faster if she is to get there on time. She is approaching a speed camera at 60mph. If she is caught speeding by the camera she will receive an £100 fine. It seems reasonable to say that if Sarah knows that the speed limit is 70mph, she is in a good enough position to go ahead and accelerate. Were her passenger to later criticise her for accelerating, she would be well within her rights to defend her behaviour by pointing out that she knew that the speed limit was 70mph. It would, I think, be unreasonable for her passenger to respond to this by saying 'Yes, you knew that the limit was 70mph. But nevertheless, you shouldn't have taken the risk'. Her passenger is too cautious. Here is another case: Rosanagh is working on her laptop in the library. She wants to go outside to smoke a cigarette, but she doesn't want to have to take her laptop with her, so she asks the girl sitting next to her to look after it while she's away from her desk. Rosanagh knows that this girl is trustworthy because they made this arrangement in the past, and everything went fine. It seems perfectly in order to say that if Rosanagh knows that the

[9] Brown does provided another apparent counterexample to the knowledge-is-always-sufficient claim in which the stakes are not high, but rather the consequences are very heavily weighted. Liz faces a bet with the following pay-off structure: if p is true, she gains a million millionths of a penny. If p is false, she loses £3. Brown claims that intuitively, even if Liz knows that p, she would be irrational to accept the bet. I don't share this judgement, but even if Brown is right, this case hardly imperils the claim that knowledge is sufficient for Eleanor, since Eleanor's consequences are nowhere near as heavily weighted as Liz's.

girl will look after her laptop and won't steal it, she is in a good enough position to go ahead and leave on the desk while she goes out to smoke. And this is true despite the fact that losing her laptop will be financially costly for Rosanagh (to the tune of, say, £500), and the benefit gained from not having to carry it around is very minor. It seems unreasonable to criticise Rosanagh on the grounds that knowledge isn't enough.

If it is true that as the practical stakes rise, a subject needs a stronger epistemic position with respect to p in order to be justified in acting on p, then it is also true that as the practical stakes decrease, a subject needs a less strong epistemic position with respect in order to be justified in acting on p than they would were the stakes higher. Accordingly, if knowledge is sufficient for the subjects in the above cases, then we should expect it to be sufficient for Eleanor too. Moreover, many cases where the stakes are roughly equivalent to those that Eleanor faces seem to deliver a clear-cut judgment that knowledge is sufficient. For example:

Cafe Case

Rosanagh is reading in the cafe. She wants to go outside to smoke, but she doesn't want to have to take her book with her. Rosanagh asks the girl sitting next to her if she will watch her book. Rosanagh would rather not take her book outside, but it will be a minor inconvenience if she has to.

Q: If Rosanagh knows that the girl next to her won't steal her book, is she okay to leave it? Or should she take it with her anyway? Or should she seek more evidence?

Toll Road Case

Rosanagh is driving to a friend's house. She knows a shortcut that will shave a few minutes off her journey time. There are lots of toll roads in the area that charge £5. But Rosanagh knows that the shortcut road isn't one of them.

Q: If Rosanagh knows that the shortcut road isn't a toll road, is she okay to take it? Or should she take the longer route just in case? Or should she seek more evidence?

Downloading Case

Rosanagh is thinking about downloading a music file. She wouldn't pay more than 10p for it were it not free, but she does want to listen to the song. If the file has a virus on it, her sister will charge her £5 to disinfect the computer. Rosanagh knows that the file is virus-free.

Q: If Rosanagh knows that the file is virus-free, is she okay to download it? Or should she not take the risk? Or should she seek more evidence?

In each of these cases, it seems clear that if she knows, Rosanagh is okay to act on that knowledge. She is okay to leave the book, take the shortcut, and download the file. Furthermore, it would be perfectly normal for Rosanagh to defend her behaviour in each case by citing the relevant item of knowledge as her justifier (e.g. 'I knew she was trustworthy, that's why I left it with her'; 'I know the shortcut isn't a toll road, so it's fine to take it'; 'I know the file is virus-free. That's why I'm downloading it'), and it would be unreasonable for someone to criticise her on the grounds that knowledge isn't enough. The stakes just aren't high enough to justify such caution.

Given that in each of these cases Rosanagh faces comparable stakes to Eleanor, if knowledge is enough for Rosanagh, then it should be enough for Eleanor too. It is counterintuitive to maintain that knowledge that p is not sufficient for p to justify Eleanor accepting the gamble. So the traditionalist who takes option 1 will have to defend counterintuitive verdicts about the circumstances in which a proposition is warranted enough to justify a subject in Φ-ing. Now, it may be that the theoretical benefits of the traditional view are weighty enough that, in the balance, these counterintuitive verdicts are a cost that should be borne. But they are a cost nevertheless, and, it seems to me, a fairly substantial one.

Prima facie, though, there seems to be a couple of ways for traditionalist's to resist this argument. One possible objection is that the intuitions I describe relating to this option aren't as clear-cut as I suggest. I have claimed that intuitions quite clearly support the claim that knowledge that p would be sufficient for p to justify Eleanor in accepting the gamble. I cite some cases in favour of this judgement where, I claim, the intuition is that knowledge is

sufficient, and the stakes are the same or less favourable. However, one might think that we can easily adjust the cases to pump intuitions in the opposite direction. So for example, consider Sarah and her passenger. Sarah's passenger says to her: "You shouldn't have risked it, even though you know that the limit is 70mph. It's not worth the risk when the fine is so high". Is it really so clear that the passenger is being unreasonable here? Arguably not. Insofar as it is unclear, the case for knowledge being sufficient is partially undermined. Similarly, one can imagine Rosanagh's mother chastising her upon hearing that she leaves her new laptop in the hands of strangers. Perhaps she is right to do so. The fact that we can seemingly make adjustments to the cases such that intuitions pump in the other direction arguably shows that intuitions can't be trusted when it comes to the decision Eleanor faces. Her's is not a clear-cut case where knowledge is sufficient, as it would be were she to face stakes of, say, p is true: win £1000, p is false: lose £1. Accordingly, I haven't posed a serious problem for the traditional view. My argument relies on hazy intuitions.

I think there are a couple of things that can be said in response to this objection. Firstly, even if it is true that intuitions are unclear in the first two cases, the potential loss for Eleanor is much lower than it is in these cases. Were a friend to criticise Eleanor's decision to accept the gamble by saying something like "I know that you know that p, but you shouldn't have taken the risk" it seems quite clear the friend is being overcautious. Eleanor would be well within her rights to defend her decision both by citing her relevant knowledge, and also by pointing out that there wasn't that much at risk - £5 isn't a lot of money. Furthermore, the three cases with comparable stakes to Eleanor's do seem like clear-cut cases where knowledge is sufficient.

Secondly, there is empirical work supporting my claim that, intuitively, knowledge is sufficient for Eleanor. Angel Pinillos (2012) has run experiments testing how much evidence the folk think is needed for knowledge in a variety of stakes-shifting cases. One such case involves John, and whether he knows that his term paper has no typos. The high-stakes situation is one where John cannot get an A for his paper unless there are no typos. He needs an A on the paper to get an A in his class, and if he doesn't get an A in his class he will lose his scholarship and have to leave college, which will be devastating for himself and his family. Separate groups of test subjects were asked the following two questions (1.) How many times does John have to check his paper before he knows there are no typos? And (2.) How many times should John check his paper before he submits it? If, in the eyes of the folk, knowledge

that his paper has no typos was insufficient for John to be justified in submitting it, given his circumstances, we would expect test subjects to offer a higher number in response to (2) than in response to (1). But Pinillos found no statistically significant difference in the answers given. This suggests that the test subjects were treating knowledge as sufficient for action in John's case. Given the comparatively minor disutility of having to check his paper relative to having to leave college, John's stakes are clearly far less favourable than Eleanor's. Accordingly, one would expect knowledge to be sufficient for Eleanor to accept the gamble in the eyes of the folk. This lends support to my claim that we can take it as common ground that knowledge would be sufficient for Eleanor.

There is another way of resisting the objection that epistemic certainty, or near epistemic certainty is necessary for Eleanor to be justified in accepting the gamble. There are not many things of which we are epistemically certain. Yet, there are many behaviours humans exhibit in day-to-day life that, when they are considered as bets, have a small positive pay-off if p is true, and a large negative pay-off if p is false. For example, suppose a friend asks me to run an errand in return for £10. Running the errand involves driving my car from A to B. Before deciding whether to take on the errand, I start thinking about the unpleasant possibility of being in a fatal car crash whilst driving from A to B, and I start thinking about whether I'm in a good enough epistemic position with respect to the proposition I won't be in a fatal car crash between A and B to be justified in accepting the errand. I need to factor the potential gains against the risk involved. Now, I'm not epistemically certain of the truth of the proposition I won't be in a fatal car crash between A and B - how could I be? - and perhaps I don't even know it. But I am pretty confident of it - after all, I'm a careful driver, and fatal car crashes are relatively rare. I decide that the possibility of a fatal car crash is small enough that I shouldn't worry about it. I go ahead and accept the errand, and everything goes fine - I don't die, and I gain £10. Now, the pay-off structure of the truth/falsity of I won't be in a fatal car crash between A and B is very negatively weighted - far more so than Eleanor's gamble, on the assumption that being in a fatal car crash is far worse than losing £5. So, should I decline to run the errand because I don't know that I won't be in a crash? According to the objection we are considering - and according to the traditionalist who takes option 1 and is consistent - if Eleanor shouldn't accept the gamble even if she knows that p, I ought not run the errand either. That seems wrong. Worse, according the traditionalist taking option 1, even if I know that I won't be in a crash, that still doesn't mean I should accept the errand. I should only do it if I am epistemically certain, or nearly epistemically certain that I won't crash. This seems

very much like the wrong result. We don't criticise people who get in their cars to run errands for risking death. Even though they don't know that they won't be in fatal car crashes. The traditionalist who takes option 1 will have to claim that such everyday behaviour is normatively defective. But, there is nothing special about the errand example. It is comparable to myriad behaviours humans exhibit every day. So the traditionalist who takes option 1, and is consistent, will have to impute massive, and highly counterintuitive, error in our judgements about what it is rational to do, and what kind of warrant is needed for a possibility to not stand in the way of a given action being justified. Insofar as we want our epistemology to cohere with common sense judgments, that's not a good position for them to be in.[10]

For these reasons, I don't think traditionalists have an easy way with option 1.[11] However, there is another way they might motivate the claim that knowledge that p isn't sufficient for Eleanor. Neta (2009) presents a case where a subject knows that p, but, Neta thinks, isn't in a good enough epistemic position with respect to p to act on it because he doesn't believe that he knows that p. Neta's case involves a student who knows that p is the answer to a test question, but doesn't believe that he knows because he is struck by a temporary insecurity about the veracity of his memory. Neta claims that in this situation the student isn't in a good enough epistemic position with respect to p for it to justify him in answering 'p'. Now, it might be that the student should write 'p' as his answer anyway, given that he has nothing to lose by attempting an answer. But intuitively, according to Neta, it is not his knowledge of p itself that justifies him answering in that way. Suppose that Eleanor is in an analogous situation. She knows that p, but she is struck by a temporary insecurity about this being an item of knowledge. Arguably, in this situation, Eleanor is not in a good enough epistemic position with respect to p for it to justify her accepting the gamble. Does this point undermine my claim that knowledge is sufficient for Eleanor? Prima facie it appears to. Eleanor knows that p, but this knowledge isn't sufficient for p to justify her accepting the gamble. Therefore:

[10] One might object to what I've just said by arguing that there is something importantly different about everyday behaviour and it's background assumptions to accepting gambles, perhaps because we don't consider the possibility of fatal car crashes when deciding whether to run errands, but do consider the possible falsity of propositions we are thinking about gambling on. But we can reformulate the point so it is analogous to Eleanor's situation. In the reformulated example, there are no errands involved. My friend offers me a gamble: If I can drive from A to B without being in a fatal car crash, he will give me £10. The traditionalist taking option 1 would have to have it that, even if I *know* that I won't be in a fatal car crash, I ought not accept the gamble. Again, this seems very wrong. Arguably I don't even need to know that I won't be in a fatal car crash to be justified in accepting the gamble. This reformulation ought to deal with that objection.

[11] Note that once we recognise these consequences, taking option 1 leads to even more counterintuitive results than it first seemed.

knowledge isn't sufficient for Eleanor.

However, troubling appearances are illusory here. The claim that if Eleanor knows that p, she is justified in accepting the gamble is a claim about propositional justification for Φ-ing. By contrast, Neta's concern is one about doxastic states and the conditions under which one justifiably Φs. The claim I'm making is that if Eleanor knows that p, then there is a proposition - p - that (propositionally) justifies Eleanor accepting the gamble. That Eleanor has such propositional justification is fully compatible with the possibility of her not justifiably Φ-ing, if she Φs. To illustrate, it is useful to analogise with the case of belief. In the case of belief, one can fail to justifiably believe q, because one believes q on the basis of r, and r isn't the kind of thing that can epistemically support a belief that q. One's belief is ill founded. Yet it can nevertheless remain true that there is some proposition p, which does (propositionally) justify a belief that q. Similarly, one can (perhaps) fail to be justified in believing that q, because one believes q on the basis of p, but one doubts that one knows, or is justified in believing, p. It can nevertheless remain true that p propositionally justifies a belief that q. Carrying this over, suppose that Eleanor Φs, and that her reason for Φ-ing is p, but, as in Neta's case, she doubts that she knows that p. Then it may be that she is criticisable for Φ-ing, since she doubted that she knew that p. Does that mean that p doesn't (propositionally) justify Eleanor Φ-ing? No. It only means, at most, that she failed to meet some necessary doxastic condition on properly using p as a motivating reason for Φ-ing. But, just as in the case of belief, there is no relevant connection - of the kind a Neta style objection requires - between meeting certain doxastic conditions on properly using a proposition as a motivating reason to perform some action Φ, and that proposition (propositionally) justifying Φ-ing. It remains true that, if Eleanor knows that p, p justifies Eleanor accepting the gamble, even if she doubts that she knows that p. The Neta style objection misses the mark.[12]

It seems to me then that traditionalists will have a hard time escaping the first horn of the dilemma. Denying that knowledge would be sufficient for Eleanor requires committing oneself to counterintuitive verdicts about knowledge, and the conditions under which it can justify action.

[12] I should note that I'm not claiming that this objection is one Neta has, or would, make against the claim that knowledge is sufficient for Eleanor. His original target was a similar, but subtly different, principle espoused in Hawthorne and Stanley (2008). The reason I have included discussion of the objection it because it seems like an objection one might wish to make to my claim that knowledge is sufficient for Eleanor.

Option 2: Claim that p is warranted enough for Eleanor

The other option available to the traditionalist is to argue that, contrary to my claim, p is warranted enough for Eleanor to justify her accepting the gamble. Is it? I said earlier that Rosanagh is a reliable testifier about whether p. That was true. But like all testifiers, Rosanagh sometimes gets things wrong. It turns out that the epistemic probability of p for Eleanor on Rosanagh's testimony, is .98 (as it is for Sarah).[13] So standard Bayesian decision theory has it that the expected monetary return for Eleanor, given her stakes and epistemic probability, is -20 pence. Now, in order to answer the question of whether p is warranted enough for Eleanor to justify her accepting the gamble, we need to know a few things about her. In particular, we need to know what her preferences are, and what her level of risk-aversion is. If it turned out that, bizarrely, Eleanor wants to lose money, then it may well be that p is warranted enough to justify her accepting the gamble. However, it is open to us to make some stipulations about Eleanor. Specifically, I'm going to stipulate the following three things are true of her:

(a) Eleanor is risk-neutral when it comes to accepting or declining the gamble

(b) Eleanor's preferences are Bayesian; so she prefers to accept the gamble if doing so has positive expected monetary return, reject it if accepting it has negative expected monetary return, and is indifferent to accepting or rejecting it if accepting/rejecting it has an expected monetary return of £0.00.

(c) Each penny gained has +1 util of value for Eleanor, and each penny lost has -1 util of value for her.

You might worry that, now that I've shared this information about Eleanor, it is no longer intuitive that were she to know that p, p would be warranted enough for Eleanor to justify her accepting the gamble. I don't think this is a serious concern. If I had suddenly stipulated something unusual about Eleanor, such that she in fact preferred to lose money from the bet, or that she is massively risk averse, then there would, of course, be grounds for complaint.

[13]Remember that we are assuming fallibilism, so it shouldn't be problematic to think that Sarah can know that p with a .98 epistemic probability for p.

But the stipulations made here about Eleanor are not unusual. They describe the preferences of an average person. And as such, it is safe to suppose that they are very similar to the assumptions that we would have made about Eleanor when assessing whether knowledge that p would have rendered p sufficiently warranted to justify her in accepting the gamble. Moreover, as we saw earlier, it seems intuitive that knowledge that p makes p sufficiently warrant to justify the actions of subjects who face far less favourable stakes than Eleanor.

Given these stipulations, what follows? The first thing to note is that, given her preferences, odds, and epistemic probability for p, accepting the gamble has a negative expected utility for Eleanor of -20 utils. So, assuming that everything is in order with the expected utility principle,[14] it would be irrational for Eleanor to accept the gamble. She shouldn't do it. Now, given her preferences, if there is any item of information that would epistemically justify Eleanor accepting the gamble, were she epistemically well positioned towards it, it is the truth of p itself. So, if p is warranted enough to justify Eleanor accepting the gamble, p epistemically justifies her accepting the gamble. However, it follows from the stipulations that p epistemically justifies Eleanor accepting the gamble only if accepting it has positive expected monetary return. But accepting the gamble doesn't have positive expected monetary return for Eleanor. So p doesn't epistemically justify her accepting the gamble.[15] Given that p is warranted enough to justify her accepting the gamble only if p epistemically justifies her accepting the gamble, it follows that p isn't warranted enough to justify Eleanor accepting the gamble. Here is that reasoning in numbered-premises form:

[14] Note that I'm not saying that the expected utility principle is right for all decision cases. As long as it works in this case, that's enough.

[15] I'm assuming here that a proposition p can't epistemically justify a subject S Φ-ing if S's epistemic position with respect to p is such that, given her preferences, it is irrational for S to Φ, but were S to have a better epistemic position with respect to p, it would be rational for her to Φ. The reason I think this is true is because one of the roles propositions can play when they epistemically justify actions is as things that can be cited in defence, and repudiation, of criticisms of Φ-ing, as in the following example: person A: "You shouldn't have sold your lottery ticket because you didn't know that it would lose"; person B (in response): "You're wrong, I was justified in selling the ticket because I knew that the lottery was rigged against me". Person A criticises person B's action - her selling the lottery ticket - on the grounds that B didn't know that the ticket would lose. B defends her action by pointing out that she *was* in a good enough epistemic position with respect to the proposition 'my ticket will lose' for it to justify her selling it. By contrast, if S shouldn't Φ - and is thus open to criticism for Φ-ing - because it is irrational for her to Φ, but it is also true that were she to have a better epistemic position with respect to p, she would *not* be irrational to Φ, then she can't properly cite p (or her epistemic position with respect to p) in defence or repudiation of the criticism that she shouldn't have Φ'ed, because Φ-ing *was* irrational, given her evidence for p, and she shouldn't have done it. However, were she to have been in a good enough epistemic position with respect to p for it to justify her Φ-ing, then she could have cited this fact to repudiate the criticism. As she can't properly cite it to repudiate the criticism, she isn't in a good enough position with respect to p for it to justify her Φ-ing. No doubt there is a lot more to be said about the link between propositions-as-justifiers and rationality. That's not a discussion I can get into here. But it seems that, at least in Eleanor's case (one where consequentialism plausibly holds), the link I'm assuming is quite hard to deny.

(4) If p is warranted enough to justify Eleanor accepting the gamble, then p epistemically justifies Eleanor accepting the gamble[16]

(5) P epistemically justifies Eleanor accepting the gamble only if accepting the gamble has positive expected monetary return for Eleanor.

(6) Accepting the gamble doesn't have positive expected monetary return for Eleanor

(7) Therefore: p doesn't epistemically justify Eleanor accepting the gamble

(8) Therefore: p isn't warranted enough to justify Eleanor accepting the gamble.

The conclusion (8) is that option 2 in untenable. In order to avoid it traditionalists will have to deny (4), (5), or (6). Premise (6) is, I take it, not open to dispute, since it is widely accepted that standard Bayesianism yields accurate verdicts on expected monetary return. That leaves traditionalists with (4) and (5). There are a couple of possible ways to deny (4), but neither of them is very compelling. Firstly, one could claim that evidence for p is simply not the kind of thing that can epistemically justify a subject accepting a gamble on p, in the same way that, say, evidence that Barack Obama is the president of the United States is not the kind of thing that can justify a belief that Jupiter has moons. Traditionalists could claim that there is no appropriate connection between p and being epistemically justified in accepting a gamble on p. But it is very hard to see why anyone would think that is true. If someone is challenged as to why they accepted a gamble on p, and responds by saying that they accepted it because they know that p is true, then it seems as though they have provided a perfectly legitimate justifier of their accepting the gamble. To claim that the fact that they knew that p was irrelevant to whether they were epistemically justified in accepting the gamble is extremely odd. One would, I think, have a hard time motivating such a claim.

Another, more subtle, way of denying (4) is to concede that p is the kind of thing that can epistemically justify Eleanor accepting the gamble - that is, concede that there is an appropriate connection - but deny the entailment relation between p's being warranted enough

[16] Note that the relevant species of justification here is propositional, rather than doxastic.

to justify accepting the gamble, and p's epistemically justifying Eleanor accepting the gamble. Jonathan Ichikawa (2012) asks us to consider the following scenario: Detective Stanley knows that the burglar came through the window. That the burglar came through the window is something that speaks in favour of arresting Smith, who is infamous for burgling houses by coming through the window, and it is amongst the things that will justify Stanley arresting Smith, if he arrests him. Clearly however, simply knowing that the burglar came through the window is insufficient reason for Stanley to arrest Smith. But that isn't because he needs a better epistemic position with respect to his belief that the burglar came through the window. He's epistemically fine with respect to the proposition. The problem is that he needs more evidence of a different kind. The point is that p can be the kind of thing that speaks in favour of Φ-ing, and can be warranted enough to justify Φ-ing, yet fail to justify Φ-ing, because p isn't a sufficient reason to Φ. Traditionalists might want to try to apply this point to premise (4). The thought would be that, whilst the truth of p can be amongst the things that justify accepting a gamble on p, even if Eleanor is warranted enough with respect to p for it to justify her accepting the gamble, she nevertheless isn't, all things considered, justified in accepting the gamble, since there is a lacuna between p being warranted enough to justify her accepting the gamble, and her actually being justified in accepting the gamble.

I find it hard to see how that line of argument can work though in this particular case though. Whilst there may well be many cases where such a lacuna is present, it's hard to see why Eleanor's case would be one of them. What would explain the gap between p being warranted enough to justify her, and p justifying her? We would want some motivation for thinking that such a gap exists, and it is not apparent what this motivation could be. It seems to me, then, that it is difficult to motivate this objection to premise (4).

Traditionalists might do better to try denying (5). But it's hard to see how this could be plausibly denied either. Given her preferences and the link we are assuming between rationality and propositions-as-justifiers, a minimal constraint on Eleanor being epistemically justified in accepting the gamble is that is reasonable for her to believe that accepting it has positive expected monetary return. That it isn't true that accepting the gamble has positive expected monetary return needn't, in itself, mean that Eleanor could not reasonably believe that it is true; subjects can, I take it, have reasonable false beliefs. But reasonable false beliefs come about when the subject has some kind of misleading evidence pointing towards the beliefs truth, and blamelessly follows that evidence. There is no reason to think that Eleanor is

blamelessly misled in the case I've described, even if she erroneously believes that accepting the gamble has positive expected monetary return. Moreover, we can simply stipulate that Eleanor's credence matches her epistemic probability, and that she has no misleading evidence of the kind that would render her blameless in believing such a thing, were she to believe that. Given these stipulations, there is little motivation for the claim that Eleanor can have a false, but reasonable, belief that she is justified in accepting the gamble, and little motivation for denying premise (5). So it looks, then, like option 2 is a hard one to take. To take it requires rejecting plausible links between rationality, expected utility, and the role of propositions justifiers of action.

IV. Summing up.

I have argued that Sarah knows that p, that if Eleanor knew that p, p would be warranted enough for Eleanor to justify her accepting the gamble, and that, given her stakes, p isn't warranted enough to justify Eleanor accepting the gamble. If these claims are correct, it follows that PE is true, since the only difference between Sarah and Eleanor is how much is at stake for them with respect to p being true, yet Sarah knows that p and Eleanor does not.

The case of Sarah and Eleanor presents traditionalists who accept fallibilism with a dilemma. In order to block the conclusion that practical stakes affect knowledge, they must deny at least one of two claims. The first denial - that knowledge would be sufficient for Eleanor - commits them to imputing massive, and highly counterintuitive, error in our judgements about what kind of warrant is needed for a possibility to not stand in the way of a given action being justified. Insofar as we want our epistemology to cohere with common sense judgement, this is a serious theoretical cost. The second denial - that p isn't warranted enough to justify Eleanor - puts them into conflict with Bayesian decision theory. This too is a serious theoretical cost. It is not obvious to me which option traditionalists should take to escape the horns of this dilemma, and any option will burden them with theoretical costs. If no option

can be plausibly taken, we should accept PE. The existential argument - along with the attendant dilemma - is quite different in form to arguments for PE currently found in the literature, and the criticisms levelled against those arguments don't have force against this one. One may well raise objections to any putative universal norms linking knowledge and practical rationality, but these objections won't affect the argument given here, since it only relies on the claim that knowledge is sufficient for action in certain conditions. In that respect it relies on a weaker premise. For this reason, it seems to me that PE theorists should prefer the argument given here.

2. Environmental Luck Gettier Cases And The Metaphysical Roles Of Knowledge

I. The metaphysical roles of knowledge

Recent epistemology has largely turned away from the project of attempting a decompositional conceptual analysis of knowledge and toward investigation of the normative and metaphysical roles that knowledge might uniquely play. To this end, on the normative side, it has been argued that knowledge is, variously, the 'norm' of belief (Sutton 2007, Williamson 2013, Littlejohn 2013), assertion (Williamson 2000, et al), and practical reasoning (Hawthorne and Stanley 2008, Fantl and McGrath 2002, 2007, 2009, et al). On the metaphysical side, Williamson (2000) argues that one's total evidence[17] includes all and only those propositions that one knows to be true ('E=K'), and Hyman (1999, 2006) argues that the fact that p can be S's reason for Φ-ing iff S knows that p:

E=K: p is part of S's total evidence iff S knows that p

reasons-knowledge thesis: the fact that p can be S's reason for Φ-ing iff S knows that p

In this essay, I focus on these two claims about the metaphysical roles of knowledge, and, by raising objections to the arguments that have been put forward in their favour, argue that they have not been sufficiently well motivated. In both cases, I suggest, the positive arguments fail for the same reason: they conflict with the results delivered by environmental luck Gettier cases.

Both E=K and the reasons-knowledge thesis have been thought by some to have far reaching implications, if true. For example, Williamson argues that if E=K is true, then it is natural to think that one is fully justified in believing that p iff one knows that p (2013). This, of course,

[17] What is your 'total' evidence? Hyman describes the idea as follows: "The ... claim that one's total evidence is simply one's total knowledge, involves the idea that an individual has a stock or fund of evidence, which can guide her assessment of hypotheses. This includes every piece of evidence a person can avail herself of, without regard to any particular hypothesis she may be considering at any particular time" (2006: 891-892). Hereafter, when I refer simply to 'evidence', I mean total evidence, unless otherwise indicated.

flies in the face of orthodoxy about the nature of justified belief, according to which it is compatible with the falsity of the proposition believed. In addition, if E=K is true, then a kind of infallibilism about knowledge can be motivated, for if one knows that p, then p is part of one's evidence, and so the probability of p on one's evidence is 1 (Williamson 2000). With respect to the reasons-knowledge thesis, Hyman (1999, 2006) argues that it motivates the thought that knowledge is best understood as a kind of ability: to know that p is to have the ability to ϕ for the reason that p. In addition, one might think that it could be marshalled in support of the claim that knowledge is the norm of practical reasoning. It is natural to think that if the fact that p can be one's reason for action iff one knows that p, then one should *treat* p as a reason for action iff one knows that p. And this is essentially what the knowledge norm of practical reasoning states. The knowledge norm of practical reasoning has, in turn, been employed to argue that knowledge is pragmatically encroached upon.[18]

Moreover, it might be thought that there are important interrelations between E=K and the reasons-knowledge thesis. In particular, that the necessity direction of the latter entails the necessity direction of the former. For, granted the (admittedly controversial) thesis that facts are simply true propositions, there are reasons to think that if the fact that p can be one's reason for Φ-ing only if one knows that p, then a proposition can be part of one's evidence only if one knows it to be true. Why? Because if p is part of S's total evidence, then it ought to be the case that S is able to believe things on the basis of p. The denial of this - the claim that there is evidence that a subject can possess, yet not employ as a basis for further beliefs - is very odd indeed, since evidence seems to be *precisely* what rational subjects form their beliefs on. But S's believing that q *on the basis of p* is simply a case of S's believing that q *for the reason that p*. It follows, if one can believe that q for the reason that p only if one knows that p, that one can possess p as part of one's total evidence only if one knows that p. And this is the necessity direction claim of E=K. The argument just given can be expressed as follows:

(1) If p is part of S's evidence, then S can believe that q on the basis of p.

(2) If S can believe that q on the basis of p, then S can believe that q for the

[18] For canonical defences of various version of the claim that knowledge is the norm of practical reasoning, see Fantl & McGrath (2002, 2007, 2009), Hawthorne (2004), Stanley (2005), and Hawthorne and Stanley (2008). Hawthorne and Stanley (2008), Brown (2008) and Neta (2009) gesture at the idea that the reasons-knowledge thesis might be marshaled in support of the knowledge norm of practical reasoning. None, however, outright endorses such an argument. Note that I am not endorsing any of the proposed implications of either the reasons-knowledge thesis, or E=K here. Rather I am merely observing that they have been made.

reason that p

(3) Therefore: If p is part of S's evidence, then S can believe that q for the reason that p

(4) By the reasons-knowledge thesis, S can believe that q for the reason that p only if S knows that p

(5) Therefore: If p is part of S's evidence, then S knows that p. (If E, then K)

As mentioned, the goal here is, by raising objections to the arguments that have been put forward in their favour, to argue that neither E=K, nor the reasons-knowledge thesis, is sufficiently well motivated. In both cases, my focus is on the (more interesting, and controversial) necessity directions of the biconditionals. That is, the claim that p is part of S's total evidence *only if* S knows that p, and the claim that the fact that p can be S's reason for Φ-ing *only if* S knows that p. I will have nothing to say about the sufficiency directions of the theses. We begin by looking at the reasons-knowledge thesis, and then move on to E=K.

II. Reasons & Knowledge

We can draw a distinction between two types of Gettier cases. 'Environmental luck' Gettier cases are cases in which one's belief that p is caused, at least in part, by the fact that p, but one fails to know that p because one is in an environment where one could easily have falsely believed p, or a similar proposition p*. 'Intervening luck' Gettier cases are cases where one's belief is causally unconnected to the fact that p, but luck intervenes to ensure that one's belief is nevertheless true. The former type, I will suggest, causes trouble for the reasons-knowledge thesis:

reasons-knowledge thesis*:* The fact that p can be S's reason for ϕ-ing iff S knows that p

The above formulation of the thesis is taken from Hyman (1999, 2006). The thesis is central to his claim that knowledge is best understood as a kind of ability. The thesis is also endorsed by Williamson (2000), Hornsby (2007), and Littlejohn (2013). Neta (2009) and Hawthorne and Stanley (2008) express sympathy to it, and Unger (1975) appeals to it to argue that nobody has a reason for doing, thinking, or feeling anything, since, he argues, nobody knows anything. I'll first sketch Hyman's motivation for the thesis, then I'll further expound on the distinction between environmental luck Gettier cases and intervening luck Gettier cases. I'll then argue that environmental luck Gettier cases undermine the thesis.[19] To be clear, my ambition here is only to show that the reasons-knowledge thesis is undermotivated by Hyman's considerations. I do not suggest that the considerations I provide furnish us with a knock-down argument against the view. Nor do I propose a rival epistemology of acting-for-reasons. Indeed, I see no *prima facie* reason to expect that there exists a true biconditional linking the ability to ϕ for the reason that p and any unique epistemic state, if the reasons-knowledge thesis is false.

2.1. Motivating the reasons-knowledge thesis

In order to motivate the necessity direction of the reasons-knowledge thesis, Hyman (1999, 2006) asks us to consider an example of a justified true belief that falls short of knowledge. Henry's belief is Gettiered:

> *Henry is watching the television on a June afternoon. It is Wimbledon men's finals day, and the television shows McEnroe beating Connors; the score is two sets to love and match point to McEnroe in the third. McEnroe wins the point. Henry believes justifiably that:*

[19] I'll assume along with these authors that facts can be reasons for the purposes of this essay. Though this is, of course, not uncontroversial.

1. I have just seen McEnroe win this year's Wimbledon final

and reasonably infers that:

2. McEnroe is this year's Wimbledon champion

Actually, however, the cameras at Wimbledon have ceased to function, and the television is showing a recording of last year's match. But while it does so McEnroe is in the process of repeating last year's slaughter. So Henry's belief (2) is true, and surely he is justified in believing (2). But we would hardly allow that Henry knows (2). (1999: 447)

Hyman then asks us to suppose that Henry recalls that his brother placed a £100 bet on McEnroe winning, and so infers that his brother has won £100. Is the fact that McEnroe is this year's champion Henry's reason for believing that his brother has won £100? Clearly not, thinks Hyman. His reason, Hyman claims, is only that he *believes* that McEnroe is this year's champion. Similarly, the fact that McEnroe won cannot be Henry's reason for being happy, angry, smug, or for congratulating his brother. Hyman takes cases like this to show that p can be one's reason for ϕ-ing only if one *knows* that p. Whenever one bears an epistemic relation to p weaker than knowledge - justified true belief, unjustified true belief, justified false belief, unjustified false belief - one's reason for ϕ-ing is not p itself, but something else. Similar considerations motivate the other authors who subscribe to the reasons-knowledge thesis (at least those who try to motivate it, rather than just endorse it - I'm thinking in particular of Unger (1975) and Hornsby (2007)).[20] Hyman goes on to argue that, given the reasons-knowledge thesis, knowledge is best conceived as a kind of ability: the ability to ϕ for the reason that p.

Although it hasn't been remarked upon before, the point extends to other cases of non-knowledge-constituting justified true beliefs that don't have the kind of double-luck structure found in Gettier cases.[21] Consider the following:

[20] Unger (1975) also appeals to linguistic data about the felicity of assertions to argue for the thesis. I won't discuss this here; though see Littlejohn (2012) for critical commentary.

[21] The double luck is this: the subject is *unlucky* to be in a scenario that is not conducive to knowledge

Sarah holds a ticket in a 1 million ticket lottery where only one ticket will win, and every ticket has an equal chance of winning. The draw has been made, and Sarah's ticket did not win, but Sarah has not seen the results yet. On the basis of the probabilities involved, Sarah believes that her ticket is a loser.

Many epistemologists have found it natural to say that Sarah's belief that her ticket is a loser is justified when based on the probabilities involved, but not knowledge. (Kyburg (1970), Klein (1985), Foley (1993), Lewis (1996), Hawthorne (2004), Pritchard (2007) et al).[22] Insofar as it is, Sarah has a justified true belief that her ticket is a loser, but does not know this. Now suppose that upon forming her belief that her ticket is a loser, Sarah becomes disappointed. Is the fact that her ticket is a loser Sarah's reason for being disappointed? Surely it is not.

2.2. Two types of Gettier case

Hyman's McEnroe case is a Gettier case where the subject's belief that p is entirely causally unrelated to the fact that p. But not all Gettier cases have this feature. There are cases of Gettiered justified true belief where the fact that p is among the causes of the subject's belief. Goldman's (1976) *Fake Barns* case is of this kind:

Henry is driving through the countryside looking at objects off the road. He sees what looks to him exactly like a barn and forms the belief 'that is a barn'. In fact it is a barn, but unbeknownst to Henry, he is in 'fake barns' country - an area with only a few real barns and many barn facades designed to look exactly like real barns to passing motorists. Luckily for Henry, he happens to look at one of the only real barns in the area. Henry's belief is justified and true, but not an item of knowledge.

---

acquisition, but then *lucky* to have their belief turn out to be true. Zagzebski (1994) was, to my knowledge, the first to notice that Gettier cases have this structure. The lottery case differs from typical Gettier cases in that intuitively it is not a matter of luck that Sarah's belief that her ticket is a loser turns out to be true.

[22] This is controversial though. See (Sutton (2007), Smithies (2012b), Smith (2010)) for the dissenting view that such lottery beliefs are not justified. I'm sympathetic to the claim that lottery beliefs are not justified, but I won't take a stand on this issue here. See the fifth essay in this dissertation 'From Moore's Paradox To The Knowledge Norm Of Belief And Beyond' for more discussion. For now I'll go along with the idea that lottery beliefs are justified.

What's the difference between the two cases? In *Fake Barns*, the fact that there is a barn is (along with various facts about Henry) one of the causes of Henry's belief that the object he saw is a barn. It figures in a causal explanation of why Henry believes what he does. By contrast, in Hyman's McEnroe case, the fact that McEnroe is this year's champion is *not* amongst the causes of Henry's belief that he is. It has no place in a causal explanation of why Henry believes what he does. Many Gettier cases have this causal-disconnect feature. Consider Chisholm's (1966) *sheep in the field* case. You are standing outside a field looking in. You see what looks to you exactly like a sheep, but is in fact a dog cleverly disguised as a sheep. You form the belief 'there is a sheep in the field'. Actually there is a sheep in the field. But it is hidden from view to you behind a hill. Your belief is both justified and true, but not knowledge. Like the McEnroe case, but unlike *Fake Barns*, the fact that there is a sheep in the field is entirely causally unrelated to your belief that there is a sheep in the field. Gettier cases come in two varieties: causal and non-causal. In the former, your belief that p is caused, at least in part, by the fact that p, but you fail to know that p because you are in an environment where you could easily have falsely believed p, or a similar proposition p*. These are 'environmental luck' Gettier cases. In the latter, your belief is causally unconnected to the fact that p, but luck intervenes to ensure that your belief is nevertheless true. These are 'intervening luck' Gettier cases.[23]

2.3. Counterexamples to the reasons-knowledge thesis

I agree with Hyman that in the case he describes the fact McEnroe is this year's champion can't be Henry's reason for ϕ-ing. But I reject the reasons-knowledge thesis because I think it is counterexampled by environmental luck Gettier cases. In order to see how such cases cause problems for the knowledge-reasons thesis, we can adapt the McEnroe case as follows:

> *Henry is watching the television on a June afternoon. It is Wimbledon men's finals day, and the television shows McEnroe beating Connors; the score is two sets to love and match point to McEnroe in the third. McEnroe wins the point. Henry*

[23] For more on the distinction between environmental luck Gettier cases and intervening luck Gettier cases, see Pritchard (2008)

believes justifiably that:

1. I have just seen McEnroe win this year's Wimbledon final

and reasonably infers that:

2. McEnroe is this year's Wimbledon champion

Henry really is watching this year's final. However, due to a network error, all other TVs in the country, including the TVs in Henry's kitchen and bedroom, and the other TV in his lounge just next to the one he is watching, are showing the final from two years ago, in which Connors staged an incredible last-ditch comeback to win the match. Only the TV that Henry is watching has been spared from the error. Henry justifiably and truly believes that McEnroe is this year's Wimbledon champion, but he doesn't know this, because he could have very easily falsely believed that Connors is.[24]

Here Henry's belief that McEnroe is this year's champion is caused, in part, by the fact that McEnroe is this year's champion. As before, suppose that Henry recalls that his brother placed a £100 bet on McEnroe winning, and so infers that his brother has won £100. Is the fact that McEnroe is this year's champion Henry's reason for believing that his brother has won £100? Unlike in the previous case, it seems obvious to me that it *is*. Similarly, were Henry to be happy, angry, or smug, or congratulate his brother, it seems clear to me that the fact that McEnroe is this year's champion would be Henry's reason for being happy, angry, or smug, or for congratulating his brother. If this is right, then, since Henry does not know that McEnroe is this year's champion, we have a counterexample to the reasons-knowledge thesis.[25] To be sure, this is a straightforward appeal to intuition. But the same is true of Hyman's argument for the reasons-knowledge thesis. And since intuition appears not to unequivocally support the

[24] This is a somewhat non-standard type of Gettier case insofar as Henry could not have easily falsely believed the proposition 'McEnroe is this year's Wimbledon champion', but rather could have easily had a false belief concerning the answer to the question 'Who is this year's Wimbledon champion?'. In this respect it more closely resembles Gilbert Harman's (1973) 'Dead dictator' case than normal Gettier cases. I take it, however, that we will nevertheless want to deny knowledge to Henry here, despite the fact that he has a justified true belief. If the reader has concerns about appeals to this kind of case, they should note that the next case that I'm about to discuss does not have this non-standard structure. It is a straightforward *Fake Barns*-style case.

[25] To be clear, the claim here is that in this case the fact that p can not only be an *explanatory* reason for Henry's ϕ-ing, but also his *motivating* reason.

thesis, we should be reluctant to accept it.

The intuitive verdict extends, I think, to many other causal Gettier cases. Consider, for example, the following variation on the *Fake Barns* case:

> *Henry is out hiking. He's lost, and the weather is turning nasty. The situation is getting serious. He sees what he believes to be a hiker's hut in the distance, and feels relieved. In fact, unbeknownst to Henry, he is in fake hiker's-hut county - an area where there are only a handful of real huts, and many hut-facades designed to look exactly like real huts to passing hikers. Henry justifiably and truly believes that the structure in the distance is a hut, but he does not know this.*

Is the fact that there is a hut in the distance Henry's reason for being relieved? Intuitively, I submit, it is. If that's right, then the reasons-knowledge thesis is false.[26]

2.4. Summing up

Certain kinds of Gettier cases - those where the fact that p causes the subject's belief that p, and so the luck preventing knowledge is environmental in kind, rather than intervening - are in tension with the reasons-knowledge thesis, and so give us cause to be reluctant in accepting it. This is perhaps not surprising, since locutions of the form 'his reason for ϕ-ing was p' can usually be replaced *salve veritate* with 'he ϕ-ed because p'. Such considerations suggest that the subject's lack of knowledge is epiphenomena in the kinds of cases that have been appealed to motivate the reasons-knowledge thesis. What was really driving our intuitions was the lack of a causal connection between the subject's belief that p and the fact that p. Does this mean that we can truthfully replace the reasons-knowledge thesis with a different biconditional

[26] I should note that some epistemologists - e.g. Hetherington (1998), Lycan (2006) - have rejected the received view that *Fake Barns* style cases are non-knowledge cases. I don't share their judgment, but I won't get into the issue here. I think it's fair to say that the view that Fake *Barns* cases are knowledge cases is a minority view. But if you hold it, then you should take the lesson of this essay to be, not that the reasons-knowledge thesis is false, but rather that it is problematic to jointly endorse *both* the reasons-knowledge thesis *and* the view that *Fake Barns* style cases are not knowledge cases. If you are particularly drawn to the reasons-knowledge thesis you might even think that the fact that *Fake Barns* cases are incompatible with the reasons-knowledge thesis gives us some reason to think that such cases *are* cases of knowledge. Similar remarks apply to the argument against E=K in the next section.

linking an epistemic state with the ability to ϕ for the reason that p? Perhaps one with a causal condition? Maybe. Maybe not. There is no reason to, it seems to me, simply *assume* that such a true biconditional exists. But the matter ought to be investigated further. That project is not taken up here, however, for even without a rival account we have good grounds on which to be hesitant in accepting the claim that the fact that p can be one's reason for Φ-ing only if one knows that p, and so hesitant in accepting the reasons-knowledge thesis. Of course, it would be hasty to take one set of cases to provide a knock-down argument against the view. Intuitions about cases can, I take it, be trumped by theoretical conditions. Nevertheless, environmental luck Gettier cases do, it seems to me, suffice to show that Hyman's argument is not persuasive.

III. Evidence & Knowledge

3.1. E=K

In the last section I argued that environmental luck Gettier cases undermine the positive case for the necessity direction of the reasons-knowledge thesis. In this section, I argue that the same is true of Williamson's arguments for the necessity direction of E=K.

Williamson (2000) argues that all evidence is propositional, and that all and only those propositions one knows to be true are part of one's evidence. Schematically, the argument has the following form:

(1) All evidence is propositional
(2) All propositional evidence is knowledge
(3) All knowledge is evidence

(C) Therefore, all and only knowledge is evidence. (E=K)

Each premise is defended with further arguments. Central to Williamson's case for E=K is the claim that one's evidence is that with which hypotheses are consistent or inconsistent. Williamson appeals to this idea in several places in the course of his argument. We first see it at work in defence of premise (1). Only propositions, it is argued, are consistent and inconsistent with hypotheses. It follows, if evidence is that with which hypotheses are consistent or inconsistent, that only propositions can be evidence. We also find Williamson appealing to the claim to argue that only *true* propositions are evidence. If one's evidence included falsehoods it would rule out some truths by being inconsistent with them. But, Williamson argues, our evidence should not outright exclude any truths, even if it may make some truths improbable. Thus, our evidence must consist of only true propositions.

My concern here is not with these arguments,[27] but with Williamson's appeal to the claim in support of premise (2), which, in conjunction with (1), expresses the necessity direction of E=K. Williamson offers two arguments in support of the claim that all evidence is knowledge. Call these the 'consistency argument' and the 'chain reaction argument'. We will look at them in turn. The consistency argument proceeds from the following case:

> *Watching a video you see a number of balls drawn from a bag in succession. Each one is replaced in the bag before the next one is drawn. You have seen draws 1 to n (for some suitable value of n); in each case the ball was red. Draw n+1 has been made but you haven't seen the colour of the ball. By reasoning probabilistically, you form the belief that the ball drawn was red. Your belief is both true and justified, but you don't know that the ball drawn was red.* (Adapted from Williamson (2000: 200))

Now Williamson asks us to consider whether either of the following two false hypotheses is consistent with your evidence at this point:

[27] E=K has attracted its fair share of critics. Here's a sample: Brueckner (2005), Silins (2005), Dodd (2007), Conee & Feldman (2008), Neta (2008), Goldman (2009), Comesana & Kantin (2010), Dancy (2011). Comesana & Kantin (2010) argue that if E=K is incompatible with the existence of a certain kind of Gettier cases. Since these Gettier cases exist, they conclude that E=K is false. My argument will also involve Gettier cases, but in a different way.

h: Draws 1 to n were red; draw n+1 was black

*h**: Draw 1 was black; draws 2 to n+1 were red.

As he points out, it is natural to say that *h* is consistent with your evidence in the case as described, and *h** inconsistent with it. More specifically, it seems to be perfectly consistent with your evidence in this case that draw n+1 was black. But if that's right, then the proposition <draw n+1 was red> cannot be part of your evidence, for then the proposition <draw n+1 was black> would be inconsistent with your evidence. By hypothesis you have a justified true belief that draw n+1 was red. So, Williamson infers, having a justified true belief that p is not sufficient for having p as part of your evidence. Williamson takes this to show that what is needed for evidence is *knowledge* that p. And this is the claim of premise (2), which is the necessity direction of E=K.

In order to assess this argument, we need to be careful in interpreting it. Textually, it is unclear exactly how Williamson intends the argument to be taken. On one possible interpretation, he takes it to apply to *all* non-knowledge-constituting justified true beliefs (hereafter 'non-K JTBs'). That is, he thinks that *whenever* one has a non-K JTB that p, it will be natural to describe ~p as consistent with one's evidence Then, with the help of further (implicit) assumptions, he deductively infers that only knowledge is evidence.

I think that it's uncharitable to interpret Williamson as holding that the argument applies to all non-K JTBs, since he doesn't outright say this. Rather, he says that an "obvious" answer to the question of why you don't have <draw n+1 was red> as evidence is that you don't know that draw n+1 was red. So a more plausible interpretation is that an abductive argument is being put forward. On this interpretation, Williamson is arguing that the *best explanation* of the intuitions elicited by the consistency argument is that only knowledge is evidence. When this is combined with the claim - in premise (3) - that all knowledge is evidence, it may well be that Williamson can plausibly maintain that E=K offers the best (simplest, most natural, most elegant etc.) account of the nature of evidence on the market.

Should we be persuaded by this argument? Williamson only considers one case. So the suggestion that E=K best explains the data seems hasty. One might well wonder whether *all* the data about when it is natural to say that a hypothesis is consistent or inconsistent with your evidence agrees with E=K. Does it? In the next section I'll argue that, once we consider more

cases - specifically, environmental luck Gettier cases - we will see that the very considerations that Williamson takes to support E=K in fact provide positive reasons to think that the contents of some non-K JTBs *are* part of one's evidence. E=K *actively conflicts* with intuitions about when a hypothesis is consistent or inconsistent with one's evidence.[28]

The chain reaction argument is gestured at in the following passage:

> If evidence required only justified true belief, or some other good cognitive status short of knowledge, then a critical mass of evidence could set off a kind of chain reaction. Our known evidence justifies belief in various true hypotheses; they would count as evidence too, so this larger evidence set would justify belief in still more true hypotheses, which would in turn count as further evidence . . . The result would be very different from our present conception of evidence. (2000: 201)

This is rather vague. How exactly would the resulting conception of evidence be very different from our present conception? Williamson doesn't say. It seems plausible, however, to interpret the argument as follows. If non-K JTBs are part of one's evidence, then in, for example, the balls-in-the-bag case, one's inferences will get inductively stronger as one continues to make inferences about draws beyond draw *n*. When you inferred that draw n+1 was red, your evidence included only the deliverances of draws 1 to *n*. However, if having a non-K JTB that p is sufficient for having p as part of your evidence, then you now have a stronger inductive base on which to infer that draw n+2 was red - namely, the fact that draws 1 to n+1 were red. This principle iterates. So, supposing that 10,000 more draws were made and all were red, if non-K JTB were sufficient for evidence you could have stronger evidence for the claim that draw n+10,000 was red than you did for the claim that draw n+1 was red. But intuitively exactly the opposite is true. Intuitively, the further away you get from your visual evidence, and the more reliant you become on induction, the *weaker* your evidence becomes that the next draw in the sequence is red.

[28] As an aside, it is worth additionally noting that if this is correct it also serves to refute the argument on its deductive interpretation (uncharitable though it is). For if we can show that considerations about when it is natural to say that a hypothesis is consistent or inconsistent with one's evidence positively support the claim that the contents of some non-K JTBs are evidence, then we will have shown *a fortiori* that there are counterexamples to the claim that, for the content of every belief that falls short of knowledge, it is natural to say that the negation of that content is consistent with one's evidence.

3.2. Against E=K

As I said, I think that environmental luck Gettier cases show that both of these arguments are unpersuasive. In order to see this we need reflect a little on Williamson's own case. Now, the case as Williamson describes it is somewhat puzzling. As several commentators have observed, unless we are sceptics about inductive knowledge, it is quite hard to see just why you are unable to know that draw n+1 was red, without actually watching the draw. Provided that n has a suitably high value, your inference that draw n+1 was red would seem to be a routine case of inductive knowledge. But if you *are* able to know that draw n+1 was red, then the consistency argument refutes E=K by Williamson's own lights, since it would refute the claim of premise (3) that all knowledge is evidence. Thus a form of inductive scepticism threatens for Williamson[29]. Nevertheless, since my interest is in premise (2), I propose to put this worry aside and grant the assumption that you have a non-K JTB that draw n+1 was red. Once this assumption is granted, the question of just *why* is it that you don't know that draw n+1 was red arises. The most familiar kind of non-K JTBs to epistemologists are those had in Gettier cases. But notice that Williamson's case doesn't look like a typical Gettier case.[30] To see this, first note that a prominent feature of Gettier cases is, as we have already seen, that the subject's epistemic environment is such that it is only a matter of luck that they believe truly.[31] But intuitively it is *not* a matter of luck in this case that your belief that draw n+1 was red is true. In fact, if anything it seems as though you would have been *unlucky* were your belief to have turned out false. Secondly, note that subjects in Gettier cases are typically not in a position to know that they don't know that p. But nothing in Williamson's case as he describes it suggests that you are not in a position to know that you don't know that draw n+1 was red. On the contrary, insofar as it is plausible that you don't know that draw n+1 was red, this seems to be something you could easily know. Thirdly, there is typically a kind of abnormality to a Gettiered subject's epistemic environment of which they are unaware, such that were they aware of it, they would no longer be justified in believing p. By contrast, there is nothing in the case as Williamson describes it suggesting any hidden abnormality. Nothing in the case suggests that you are not aware of all the relevant facts about your environment. Insofar as you can justifiably believe that draw n+1 was red, it seems that you can do so in

[29] This point has been made by Dodd (2007), Weatherson (manuscript), and McGlynn (2014)
[30] McGlynn (2014) also makes this observation.
[31] Pritchard (2012) describes this as one of the 'master intuitions' driving Gettier cases.

full awareness of all the relevant facts about the case.

In short, Williamson's case lacks many of the hallmarks of Gettier cases[32] [33]. Given this, one may well wonder if the consistency argument is persuasive when run against a case that does have these hallmark features. I don't think it is. To see this, consider the following variation on Williamson's case, where you *do* see draw n+1:

> *Watching a video you see a number of balls drawn from a bag in succession. Each one is replaced in the bag before the next one is drawn. You have seen draws 1 to n (for some suitable value of n); in each case the ball was red. You then see draw n+1, in which again a red ball is drawn from the bag. Unbeknownst to you however, between draws n and n+1 the bag was surreptitiously switched. The new bag, from which draw n+1 was made, contained hundreds of black balls disguised as red balls, and one genuinely red ball.*[34] *By sheer luck, the genuinely red ball was drawn at n+1. You justifiably and truly believe that draw in n+1 was red, but you don't know this because your belief is Gettiered - you would have believed that draw n+1 was a red ball even if it was really a disguised black ball.*

In this case you don't know that draw n+1 was red, because your belief is Gettiered. Since your belief was caused, in part, by the fact that draw n+1 was red, you are in an environmental luck Gettier case. Now consider the following false hypotheses:

> *h:* draw n+1 was black.

---

[32] Plausibly, the case is much better understood as analogous to a lottery case - a case where you believe solely on the basis of the probabilities involved that your lottery ticket is a loser. This raises interesting issues in itself, since a number of epistemologists (e.g. Smithies (2012), Smith (2010), Sutton(2007), amongst others) have argues that you not only cannot know that your ticket has lost solely on the basis of the probabilities involved, but you cannot even justifiably outright believe that it has. In that case, we might wonder if Williamson's description of the case as involving a *justified* belief is correct. Regrettably, I cannot go into these issues here. But the important thing to note is that the case bears little resemblance to a typical Gettier case

[33] The phrase 'Gettier case' is sometimes used to refer to any case of a justified true belief that is not knowledge, irrespective of the specific features of the believers epistemic situation. I do not use the phrase in that way. I'm taking 'typical Gettier case' to refer to cases that have the features outlined above. If the reader is unhappy with this, they should mentally replace the phrase 'not a Gettier case' with 'not a case where the subject is lucky that they truly belief that p, nor in a position to know that they don't know that p, nor in an abnormal epistemic environment such that were they aware of the abnormality, they would no longer be justified in believing that p'.

[34] Of course, it might be argued that a black ball disguised as a red ball simply *is* a red ball. I ask the reader to put this concern aside. The case could be easily amended to get around the worry. But I prefer to keep it as similar to Williamson's case as possible.

If consistency considerations favour E=K, it should be natural to say that *h* is consistent with your evidence in this new case. Is it? I feel no inclination whatsoever to say that it is, and to my ears it would not be natural to describe it as such. This point alone is troubling for Williamson's argument. But we can go further. Insofar as considerations about when it is natural to say that a hypothesis is consistent with one's evidence are a good guide to what propositions are *not* in one's evidence - and of course Williamson's argument trades on the assumption that they are - then considerations about when it is natural to say that a hypothesis is *inconsistent* with one's evidence must also be a good guide to what propositions *are* in one's evidence. And it seems to me that not only would it not be natural to describe *h* as consistent with your evidence in this case, it would be natural to describe *h* as positively *inconsistent* with your evidence. If that's right, then by Williamson's own endorsed method for establishing just which propositions are and are not part of one's evidence, we get the result that the contents of some non-K JTBs *are* evidence. That is, we get, by Williamson's own method, positive reasons to think that some of these contents are evidence. Thus, if we are to accept the method of the consistency argument at all, we must accept that it gives us positive reasons to think that the contents of some non-K JTBs are evidence, and thus positive reasons to think that E does not equal K.

The verdict of the above case is not a one-off. New cases could be created in which the same intuition is elicited.[35] Although there is no scope for a full diagnosis in this paper, it seems to me that all environmental luck Gettier are cases where it is natural to say that the negation of the subject's belief is inconsistent with their evidence. Note that this is not to say that in *all* Gettier cases it is natural to say that the negation of the subject's belief is inconsistent with their evidence. As we have seen, Gettier cases also come in the intervening luck variety, and it is less clear that such cases conflict with Williamson's claims about consistency with the evidence.

Williamson's consistency argument fails to support his conclusion. Whilst we may accept a view on the nature of evidence on the basis of an abductive argument that is silent on some cases, we should be reluctant to accept the view on the basis of an argument that turns out to issue results that *actively conflict* with the view, even if we don't yet have a rival explanation

[35] Littlejohn (2012) argues that Goldman's fake barns case, which is structurally similar to mine, causes problems for E=K. Littlejohn's argument, however, trades directly on the intuition that the driver has the same evidence when they are driving in real-barn country and fake-barn country, rather than engaging with Williamson's consistency argument.

of the data.

What about the chain reaction argument? Is this more persuasive than the consistency argument? It is not, for it too fails to show that non-K JTBs of the kind found in environmental Gettier cases (or, for that matter, intervening luck Gettier cases) cannot be part of your evidence. Insofar as the argument is persuasive, all it shows is that non-K JTBs *arrived at by induction or probabilistic reasoning* cannot be part of your evidence. But of course, it is possible to have a non-K JTB that p that isn't arrived at by induction or probabilistic reasoning, but rather by *observation*, as is the case in the environmental luck version of the balls-in-the-bag case. And when your belief is arrived at in such a way, no problematic chain reaction of the kind described is set off. Thus, again, Williamson's argument falls short of supporting his desired conclusion.

Of course, as with the argument against Hyman, these aren't knock-down arguments against Williamson's position; they don't entail that E=K isn't ultimately the right view to adopt. Perhaps, for example, rival views that better accommodate the data issued from consistency considerations will falter for different reasons. Or perhaps such views will match E=K in accounting for a wide range of data, but lose out on other theoretical virtues such as simplicity, elegance, and naturalness. Nevertheless, it’s clear that the consistency and chain reaction arguments don't motivate E=K. The chain reaction simply does not support the thesis, and if anything, consistency considerations motivate E≠K. As such, Williamson's arguments that only knowledge is evidence fall short of motivating the claim.

IV. Conclusion

I have argued that environmental luck Gettier cases cause problems for arguments purporting to show that knowledge has two important metaphysical roles: that one's evidence consists of

all and only those propositions that one knows to be true, and that the fact that p can be one's reason for Φ-ing iff one knows that p. I have, however, been careful not to claim that such cases provide us with knock-down arguments against the respective views. Nevertheless, more work will need to be done in order to make a persuasive case that knowledge does in fact play these roles.

3. Might The Simulation Heuristic Influence Knowledge Attributions?

I. Introduction

Recently epistemologists have noticed an intriguing pattern in people's knowledge attributions. People tend to be less willing to attribute knowledge-that-p to a subject when specific unrealized metaphysical possibilities of error are mentioned, and so made salient, than they are when such possibilities are not mentioned and not made salient, but present nevertheless. Call this the 'error-salience phenomenon'. Some have appealed to the phenomenon to motivate contextualist theses about the semantics of 'knows', and it appears to pose a challenge to traditional invariantist approaches to the word's semantics.

What, if any, psychological processes might explain the error-salience phenomenon? The question has only recently generated interest. John Hawthorne (2004) and Timothy Williamson (2005) speculatively suggest that the phenomenon might be the result of the influence of the *availability heuristic* (Kahneman & Tversky 1972). They appeal to this idea to undermine the force of the apparent challenge the phenomenon poses to invariantism. Jennifer Nagel (2010) rejects Hawthorne and Williamson's proposal, and instead makes the case that the phenomenon is the result of *epistemic egocentricity* (Royzman *et al* 2003, Birch *et al* 2004). Mikkel Gerken (2013) proposes an explanation in terms of *epistemic focal bias*. Like Hawthorne and Williamson, Gerken and Nagel appeal to their proposals to attempt to undermine the force of the apparent challenge that the error-salience phenomenon poses to invariantism

In this essay I make a preliminary case that the operation of the *simulation heuristic* (Kahneman & Tversky 1982) may partially explain the error-salience phenomenon. Unlike Hawthorne, Williamson, Nagel, and Gerken, my main motivation is not to defend invariantism, but first and foremost simply to understand the phenomenon in question. The project is an interesting one in its own right. If epistemologists want to fully understand the nature of knowledge, then it is imperative that they understand the psychology of knowledge attribution. Nevertheless, I will make some (largely negative) remarks about the prospects of defending semantic theses about the word 'knows' by appeal to psychological explanations of knowledge attribution towards the end of the essay.

The plan is this: §2 sets the scene, introducing the evidence - both intuitive and empirical - for the error-salience phenomenon, and briefly describes the impact it has had on debate about the semantics of 'knows'. §3.1 introduces the simulation heuristic, some of its interesting properties, and some of the data that has been adduced in favour of postulating it. §3.2 makes the case that the error-salience phenomenon can plausibly be explained by the hypothesis that knowledge attributions are influenced by the simulation heuristic. §4.1 presents Hawthorne and Williamson's availability heuristic explanation. §4.2 discusses Nagel's objections to it. §4.3 considers how the simulation heuristic explanation hypothesis fares against Nagel's objections. §4.4 presents Nagel's epistemic egocentricity account. §4.5 considers the relationship between Nagel's account and mine, and consider what might be said in favour of each. In §4.6 I show how my account integrates with Gerken's epistemic focal bias account. In §5, I discuss the implications of the proposed psychological explanations on the debate about the semantics of 'knows'. §6 concludes.

Before we proceed, a disclaimer is needed. Questions about the psychology of knowledge attribution are complex and relatively underexplored. We should not expect to arrive at well-supported conclusions about the influence of psychological mechanisms on knowledge attribution without considerable empirical investigation and theoretical reflection. I make no pretense that the proposal sketched here is definitive, nor that the argument offered in favour of it is conclusive. Nevertheless, productive inquiry is guided by the assessment of *prima facie* plausible hypotheses. The goal of this essay is to put forward just such a hypothesis. In §4.3 and §4.5 I discuss some outstanding empirical questions that need to be answered if we are to move forwards.

II. The Semantics of 'Knows'

Consider the following two vignettes, borrowed from Nagel (2010):

(a) The plain story: John is in a furniture store. He is looking at a bright red table under normal lighting conditions. He believes the table is red.

Q: Does John know that the table is red?

(b) The more detailed story: John is in a furniture store. He is looking at a bright red table under normal lighting conditions. He believes the table is red. However, a white table under red lighting would look exactly the same to him, and he has not checked whether the lighting is normal, or whether there might be a red spotlight shining on the table.

Q: Does John know that the table is red?

Many epistemologists find it easier to attribute knowledge, rather than mere true belief, to John in case (a) than in case (b),[36] despite the fact that, we may suppose, there no difference in John's epistemic situation between (a) and (b). Case (b) merely draws attention to a counterfactual aspect of John's situation that is unmentioned, but nevertheless present, in case (a). Although many of them regard it as a mistake, something approaching a consensus has emerged amongst epistemologists that we are, *as a matter of fact*, more reluctant to attribute knowledge when unrealized possibilities of error are mentioned, and so made salient, than we are when such possibilities go unmentioned. For ease of exposition, we will refer to cases like (b) as 'mentioned-error-possibility cases' (hereafter 'MEP-cases'), cases like (a) as 'non MEP-cases', and the phenomenon of interest the 'error-salience phenomenon'.

For some time the existence of the error-salience phenomenon was taken for granted by epistemologists. Recently it has been put to empirical test. Do non-philosophers share philosophers intuitions about MEP and non-MEP cases? Initial studies suggested a negative answer (Buckwalter 2010, Feltz and Zarpentine 2010). More recent studies, however, have found some evidence for the phenomenon. For example, Schaffer and Knobe (2012) presented participants in an experiment with variations on Keith DeRose's (2009) bank cases, in which Hannah asserts to her partner that the bank is open on Saturday. Some participants were given vignettes in which the possibility of error was made salient. Others were given vignettes where this possibility was not made salient. Participants were then asked to rate their degree of agreement or disagreement with the statement: 'Hannah knows that the bank

[36] A sample: Hawthorne (2004), Lewis (1996), Stanley (2005), DeRose (2009), Unger (1971), Cohen (2002).

will be open on Saturday'. Schaffer and Knobe found that participants presented with the non MEP-case were more inclined to agree with the statement than those presented with the MEP-case, to a statistically significant degree. Alexander, Gonnerman, and Waterman (forthcoming) found the same with Nagel's (a) and (b) cases. Further confirmatory work is needed, but from here on in we will assume along with the rest of the participants in the debate of interest that error-salience is a real and robust phenomenon.

The error-salience phenomenon has played an important role in recent debate about the semantics of the word 'knows'. Contextualists hold that the truth conditions of sentences of the form 'S knows that p' are context-sensitive. One popular way of cashing out this idea is that the context partially determines the epistemic standard S needs to meet in order for the sentence 'S knows that p' to be true in that context. Contextualists have claimed to find vindication for their view from the error-salience phenomenon.[37] If contextualism is true, it is perfectly consistent to claim that 'John knows that the table is red' in non MEP-cases like (a) and 'John doesn't know that the table is red' in MEP-cases like (b). The standards required for it to be true to say that 'John knows that the table is red' are higher, the thought goes, in (b) than they are in (a), because an error possibility has been made contextually relevant in (b) but not (a). DeRose (2005) is quite explicit about the error-salience phenomenon being a central motivation for contextualism:

> "The best grounds for accepting contextualism concerning knowledge attributions come from how knowledge-attributing (and knowledge-denying) sentences are used in ordinary, non-philosophical talk: What ordinary speakers will count as ‘‘knowledge’’ in some non-philosophical contexts they will deny is such in others" (2005: 172)

By contrast, the error-salience phenomenon seems, *prima facie,* to pose a challenge to invariantists who hold that 'knows' does not have a contextualist semantics. If the truth-conditions for sentences of the form 'S knows that p' aren't shifty in the way that contextualists suggest, then why are people's knowledge attributions shifty? Invariantism appears to be at odds with actual usage. And this, one might think, is a mark against it. (DeRose 2009, et al.)

Against this background, some invariantists have sort to motivate a certain kind of error-

[37] DeRose (2005, 2009), Lewis (1996), Cohen (2002)

theory about the error-salience phenomenon. Their strategy is to develop hypotheses about the psychological mechanisms that produce the phenomenon whereby it is the result of well-known heuristics and biases, and then argue that these psychological explanations of the phenomenon show that it does not undermine invariantism. So we see Hawthorne (2004) and Williamson (2005) suggesting that the phenomenon is the result of the influence of the availability heuristic, Nagel (2010) arguing that it is the result of epistemic egocentricity, and Gerken (2013) arguing that it is the result of an epistemic focal bias.

My interest here is not primarily in the debate between invariantists and contextualists (though see §5 for some remarks on the prospects of appealing to psychological explanations to defend classical invariantism). Rather I am first and foremost interested in understanding what psychological mechanisms give rise to the error-salience phenomenon. The proposal put forward here is that the influence of the simulation heuristic might at least partially explain it. In the next section I introduce the heuristic and make the case for its ability to explain the phenomenon.

III. The Simulation Heuristic Hypothesis

3.1 Introducing the simulation heuristic

Counterfactual thinking - thinking about what was not, but could have been, the case - is pervasive in human cognition. It is, amongst other things, central to the imagining of novel combinations that is a building block of creativity (Byrne 2002, Costello & Keane 2000, Fink *et al.* 1992). As philosophers well know, it also helps people to reveal counterexamples to deductions (Johnson-Laird & Byrne 2002, Byrne *et al.* 1999), and judgements of accountability, fault, responsibility, and blame, are often the consequence of counterfactual

thought (Markman & Tetlock 2000, Hegarty & Pratto 2000).

Psychologists have converged on agreement about the way that counterfactual judgement is produced. As Kahneman and Tversky (1982) put it, the assessment of counterfactuals is produced by an operation that resembles the running of a simulation model. When considering whether a proposition like 'If LeBron had made his last 3-point attempt, the Miami Heat would have won the game' is true, people will mentally simulate alternative sequences of events and their outcomes by mutating events that obtained in the actual world. For example, a very obvious mutation for assessing whether the proposition 'If LeBron had made his last 3-point attempt, the Miami Heat would have won the game' would be to imagine a world in which Lebron *actually* makes the last 3-pointer, and then running a simulation to form a judgement about what might have happened thereafter.

This kind of psychological process is heuristically governed. As Byrne (2002) puts it, people systematically 'zoom in' on the same things from potentially infinite sets of possibilities. That is, they will, with a high degree of regularity, mutate certain possibilities rather than others in forming counterfactual judgements. For example, studies show that people tend to mutate exceptional events rather than normal ones (Kahneman and Tversky 1982), tend mutate events that are in their control more than those that are out of their control, tend to mutate recent events more than temporarily distal events within causal chains (Miller & Gunasegaram 1990, Segura *et al.* 2002, Byrne *et al*. 2000), and tend not to mutate natural laws (Seelu *et al.* 1995). It is these tendencies that go under the name 'the simulation heuristic'.

3.2. Making the case for an explanation of the error-salience phenomenon from the influence of the simulation heuristic

My proposal is that the operation of the simulation heuristic might at least partially explain the error-salience phenomenon. How might this happen? Our starting point is the observation that there is wide agreement amongst epistemologists that knowing-that-p is a modal state, requiring that one's belief that p is true in a certain set of metaphysically possible, but non-

actual, worlds where it is held. There are a number of different ways to flesh out this idea. For example, the sceptically inclined might take knowing to require that one's belief is true in *all* or *almost all* possible worlds where it is held, thereby placing a very high bar on knowing. More popular are views that place restrictions on membership of the set. According to those who advocate a *safety* condition on knowledge, one's belief need not be true in all possible worlds in which it is held, but only all those possible worlds that are *nearby* to the actual world, where nearness is determined by a similarity relation. According to those who advocate a *sensitivity* condition, knowing requires that one not believe that p in the nearest worlds in which p is false.[38]

However one chooses to understand the modal requirements on knowledge, the following hypothesis is very plausible if knowing-that-p is a modal state and assuming that people are generally accurate in their knowledge attributions: their actual assessments of whether a subject, S, knows that p, will involve counterfactual thinking. This certainly seems to be the case when people are presented with Gettier cases, for example. A plausible diagnosis of why people tend to deny knowledge in Gettier cases is that they recognise that there are very nearby possible worlds in which the subject falsely believes that p, and therefore judge that the belief is not sufficiently modally robust to count as an item of knowledge. The subject, they recognize, could easily have falsely believed that p.

If people's assessments of whether a subject knows do proceed in this way, then it would, one might think, be no surprise were the simulation heuristic to influence knowledge attributions. Much counterfactual thinking is influenced by the simulation heuristic, and knowledge attribution involves counterfactual thinking, so it is *prima facie* plausible that the simulation heuristic will influence knowledge attribution. But this observation alone does not provide us with any clues as to just *how* the simulation heuristic might exert its influence, and thus does not provide us with any kind of explanation of the error-salience phenomenon. To find such an explanation, we need to look more closely at the workings of the simulation heuristic.

One key observation comes from Kahneman and Miller (1984). Kahneman and Miller observed that in performing counterfactual reasoning about what could have been the case,

[38] For defences of the claim the safety is necessary for knowledge, see Williamson (2000), Pritchard (2005) and Sosa (1999). For defences of the claim that sensitivity is necessary for knowledge, see Nozick (1981) and DeRose (2009).

people will tend to treat some aspects the situation as relatively fixed and immutable, and others as mutable. Which aspects of the situation are treated as mutable, and which are treated as immutable, is affected by the simulation heuristic. Kahneman and Miller hypothesise that the mutability of any aspect of a situation *increases* when attention is drawn to it, and it is made psychologically salient. Evidence for this hypothesis, they note, was found by Read (1985). In one of Read's experiments, subjects were taught the rules of a simple two-person card game. Subjects were then shown the hands of two players, A and B, and asked to complete stems such as, "A would have won if..." or A would have lost if...." Read found that a large majority of completions involved changes to the hand held by A, despite the fact that the same outcome could, of course, been just as easily produced by altering B's hand. The fact that subjects tended to mutate the actions of the actor to whom attention is drawn, and tended not to mutate the actions of the actor to whom attention was not drawn, supports Kahneman and Miller's hypothesis. Read also tested the hypothesis with vignettes such as the following:

> *Helen was driving to work along a three-lane road, where the middle lane is used for passing by traffic from both directions. She changed lanes to pass a slow-moving truck, and quickly realizes that she was headed directly for another car coming in the opposite direction. For a moment it looked as if a collision was inevitable. However, this did not occur.* (Read 1985)

Read asked subjects to indicate, in one sentence, how they think the accident was avoided. A substantial majority of subjects completed the story by ascribing the critical action to Helen. Again, this supports Kahneman and Miller's hypothesis that drawing attention to an aspect of the situation increases its mutability.

Questions remain about just how strongly Read's findings support Kahneman and Miller's hypothesis that the mutability of *any* aspect of a situation increases when attention is drawn to it, since Read's experiments involve the mutability of the actions of focal and background *actors*, and it doesn't follow from the fact that subjects tend to mutate the actions of actors to whom attention is drawn that they will tend to mutate *any* aspect of the situation to which attention is drawn. Nevertheless, Read's studies lend some credence to Kahneman and Miller's hypothesis. The matter needs to be investigated further, but let us suppose that Kahneman and Miller are correct in their hypothesis. What implications might this have for our understanding of the psychological processes producing the error-salience phenomenon?

A natural thought is this: if Kahneman and Miller are right, and the mutability of any aspect of a situation increases when attention is directed to it, then we should expect readers to be more reluctant to attribute knowledge in MEP' cases than in non-MEP cases. To see why this is, consider Nagel's John cases from §2. In case (a) no attention is drawn to the fact that John would have falsely believed that the table was red if he was looking at a white table under a red light. Case (a) is a non-MEP case. Case (b) is a MEP case; attention is drawn to the fact that John would have falsely believed that the table was red if he was looking at a white table under a red light. If Kahneman and Miller's hypothesis is correct, then in case (b) this fact of the situation will be highly mutable (or at least, more mutable than it is in case (a)). Readers of (b) will thus be more prone to imagine it *actually* obtaining. In doing so, they will come to recognise that there is a possible world in which John falsely believes that the table is red. By contrast, they will display less of a tendency to arrive at this conclusion in case (a). Since attention is not drawn to the fact that John would have believed that the table is red if it was white but he was looking at it under a red light, this aspect of the situation will tend not to be entertained and will tend not be mutated. Readers will thus tend not to recognise that this is a possible world in which John falsely believes that the table is red.

How the explanation proceeds from here depends on how strong the demands are that readers place on subjects in order to count them as knowing. If people demand that a subject's belief is true in all possible worlds in which it is held in order to count that subject as knowing, then we are availed of an explanation of how the influence of the simulation heuristic might produce the error-salience phenomenon. In MEP cases like (b), but not in non-MEP cases like (a), readers will be more likely to come to recognise that there is a possible world in which John falsely believes that the table in from of him is red, and will thus be more likely to deny that he knows that it is red.

However, suppose that people don't demand truth in *all* possible worlds in which the belief is held, but only in a certain epistemically relevant set of worlds. Then the denial of knowledge will depend on whether readers process the world in which John falsely believes that he is looking at a red table because he is looking at a white table under a red light as an epistemically relevant world. Why should we think that they would do that? One reason to expect such an outcome is that people tend to treat the salience of something as evidence for

its relevance to the task at hand (see, e.g., Gigerenzer and Todd 1999).[39] If so, then a plausible explanation emerges for why readers will tend to deny that subjects know in MEP cases like (b) even if they don't demand that the subject's belief is true in all possible worlds in which it is held. The world in which John falsely believes that he is looking at a red table because he is looking at a white table under a red light is contextually salient in (b). So readers will tend to process it as epistemically relevant - that is, they will tend to process it as one of the worlds in which John has to have a true belief in order for his belief in the actual world to count as knowledge. From the operation of the simulation heuristic, they will come to recognise that he would falsely believe that the table is red in this world. Since John falsely believes that the table is red in this world, and subjects will tend to treat this world as epistemically relevant, they will therefore tend to deny that he knows that the table is red in the actual world. This will not happen in the (a) case. Thus we would expect to see higher levels of knowledge attributions in (a) than in (b).

The focus here has been on Nagel's John cases. But they are merely an example. The point should generalise to all of the error-salience cases that epistemologists have been interested in, since they all share the common feature of drawing attention to certain unrealized error possibilities. In that case, we have a (preliminary) case for thinking that the error-salience phenomenon is at least partially the result of the influence of the simulation heuristic.

IV. Comparison With Rival Accounts

§4.1. The availability heuristic explanation

[39] Gerken (2013) also makes this point. Gerken's account will be discussed in more detail in §4.6

In the last section, I made the case that the simulation heuristic might at least partially explain the error-salience phenomenon. Another explanation for the phenomenon can be found in Hawthorne (2004) and Williamson (2005). Hawthorne and Williamson suggest that the phenomenon might (at least in part) be the result of the influence of the *availability heuristic*. In this section and the next I discuss the availability heuristic explanation and objections to it from Nagel (2010). The purpose of doing so is to assess how the simulation heuristic explanation fares against Nagel's objections. It is important to do so, since the simulation and availability heuristics are thought to be closely related. Thus one might worry that the problems Nagel identifies with the availability heuristic explanation will be inherited by the simulation heuristic explanation. This will be a starting point for critical evaluation of the simulation heuristic explanation.

The availability heuristic was first postulated by Kahneman and Tversky (1973). Roughly characterised, it describes a psychological phenomenon whereby people rely on the ease with which examples of events of a certain type can be imagined or remembered in order to form judgements about the probability that an event of that type will occur. Like all heuristics, the availability heuristic provides a cognitively cheap way of arriving at a conclusion with a fairly good degree of accuracy. However, the cost of employing this mental shortcut is systematic fallibility. Relying on the availability heuristic leads people to overestimate the likelihood of certain events that are uncommon, but easily imagined and remembered. For example, people tend to overestimate the number of deaths caused by dramatic events, since such events are disproportionately discussed.[40] Could something similar be going on in epistemology? Williamson thinks it's a possibility:

> "One effect of fictional violence on television is to make viewers overestimate their chances of being the victims of violent crime: they suffer an illusion of danger. Might not an illusion of epistemic danger result from exposure to lurid stories about brains in vats, evil demons, painted mules, or gamblers who bet the farm?" (2005: 226)

The suggestion - put forward by both Williamson and Hawthorne - is that the effect of making unrealized error possibilities salient in MEP-cases is to lead people, by relying on the availability heuristic, to overestimate the likelihood of their obtaining, and thereby come to

[40] Lichtenstein *et al*. (1978)

deny that subjects in MEP cases know. Hawthorne cites Slovic *et al.* on this point. Says Slovic: "[a] particularly important implication of the availability heuristic is that discussion of a low-probability hazard may increase its memorability and imaginability and hence its perceived riskiness" (Slovic *et al*. 1982: 465).[41]

§4.2. Nagel's objections to the availability heuristic explanation

It is important to note at this juncture that the simulation heuristic explanation and the availability heuristic explanation needn't be seen as rivals. Hawthorne makes it clear that he does not necessarily take the availability heuristic explanation to exhaustively account for the error-salience phenomenon. Other psychological mechanisms may also contribute. One of these, it might be thought, could be the simulation heuristic. Similarly, I am not suggesting that the simulation heuristic explanation exhaustively explains the phenomenon. It is perfectly compatible with my account that the availability heuristic also contributes. That said, there are good reasons to think the availability heuristic is *not* contributing to the phenomenon, for Williamson and Hawthorne's proposal has been subjected to criticism from Nagel (2010). In this section I present, and comment on, Nagel's objections. Given the similarities between the availability heuristic and the simulation heuristic, Nagel's objections will be presented in some detail, since the task of the next section is to discuss how the simulation heuristic explanation fares against them.

§4.2.1 The imaginability objection

Nagel's first objection departs from the work of Steven Sherman (1985). Sherman found evidence that when a given event is difficult to imagine, subjects who have been encouraged to imagine or discuss it will tend to assess its occurrence as *less* probable than subjects who have not been encouraged. Sherman presented subjects with booklets about a fictitious disease, which described symptoms of a disease without giving any statistics on its prevalence (though it was described as becoming 'increasingly prevalent' on the subject's student campus). One group was given a booklet in which the described symptoms were easy to

[41] It is important to note that Hawthorne and Williamson don't fully endorse the availability heuristic explanation. They merely put it forward as a interesting tentative possibility.

imagine. The other group was given a booklet in which the described symptoms were hard to imagine. Within each group, half of the participants were given the booklet and asked to assess the likelihood of their contracting the disease, and half were asked to imagine themselves coming down with the disease over a three-week period and then write out a narrative detailing how they imagined they would be affected. They were then asked to assess the likelihood of their contracting the disease. Sherman's results were as follows: subjects who read the hard-to-imagine booklet, but did not imagine themselves coming down with the disease, rated their probability of contracting the disease as slightly lower than their counterparts who read the easy-to-imagine booklet. Subjects who read the hard-to-imagine booklet *and* imagined themselves coming down with the disease, rated themselves *less* likely to contract the disease than their counterparts who did not engage in the imagining task. Subjects who read the easy-to-imagine booklet and engaged in the imagining task rated themselves as more likely to contract the disease than their counterparts who did not engage in the imagining task.

What Sherman's study shows, Nagel notes, is that being encouraged to imagine a scenario only increases its apparent probability when the scenario is easy to imagine. When the scenario is hard to imagine, the opposite effect can occur. This is where Nagel's objection kicks in. Nagel argues that it is far from clear that the kinds of MEP-cases to which epistemologists trade involve unrealized error possibility scenarios that are easy to imagine: "It is not immediately obvious that scenarios involving brains-in-vats and tricky lighting should count as 'easy to imagine' in the relevant sense (to say nothing of more esoteric examples involving brain lesions and strange coincidences). Casual introspection is unlikely to settle the question of whether a given scenario has a level of imaginability that would prompt a positive rather than a negative availability effect" (2010: 293). Nagel's point here is not to make the case that the availability heuristic *isn't* doing the work here. But rather that it is far from obvious that it *is*. Call this first objection the *imaginability objection*.

§4.2.2 The comparative imaginability objection

Nagel's second objection leads on from the imaginability objection. Nagel observes that even if the sceptical possibilities that epistemologists trade in are easy to imagine (and so the imaginability objection doesn't hold), we should expect to see *higher* levels of knowledge-denial in MEP-cases where the unrealized error possibility is commonplace and unexotic than MEP-cases where it is arcane, recherché, and hard to imagine. But it is not obvious, she

argues, that this is the case. If it isn't, then it is unlikely that the availability heuristic explanation can account for the error-salience phenomenon. Call this the *comparative imaginability objection*.

§4.2.3 The spontaneous discounting objection

Nagel's third objection departs from the observation that what the availability heuristic explanation needs to account for the error-salience phenomenon is that *sheer mention* of unrealized error possibilities causes people to overestimate the likelihood of the possibility via the availability heuristic. Why is sheer mention necessary? Because the phenomenon that needs to be explained is that people tend to shift in their knowledge attributions from making positive attributions in non MEP-cases like (a) to negative attributions in MEP-cases like (b) *when only MEP-cases mention the possibility of error.* Drawing on the work of Daniel Oppenheimer (2004) *et al*., and Norbert Schwarz *et al.* (1983), Nagel observes that sheer mention tends to have the *opposite* effect to what those who endorse the availability heuristic explanation need here. In a series of experiments Oppenheimer found that when subjects recognised that there was an explanation for increased availability that was unrelated to objective frequencies, they would automatically discount the significance of availability in forming judgements - a tendency known as *spontaneous discounting*. For example, in one experiment Oppenheimer asked subjects to rank the relative frequencies of various surnames in the USA. The subjects were generally quite good at this. But Oppenheimer discovered that they had a tendency to *underestimate* the objective frequency of famous names like 'Bush', 'Clinton', 'Furtado', and 'Kravitz'. Oppenheimer's suggestion was that this result is the result of spontaneous discounting: the subjects will have recognised that the increased availability of such names was because the people they associated with them were famous, recognised that the fame of a name does not correlate with its objective frequency, and so discounted availability in forming their judgements about the objective frequency of the name in the USA. Spontaneous discounting also occurs across a range of other types of judgements. For example, Schwarz *et al.* (1983) found that, whilst subjects tend to report greater happiness on sunny days than on rainy days, this phenomenon can be cancelled by explicitly drawing subject's attention to the weather by commenting on it immediately before asking them about their mood.

Supposing that the availability heuristic is affected by spontaneous discounting, how does this cause problems for the availability heuristic explanation of the error-salience phenomenon?

Nagel observes that the general lesson of spontaneous discounting is that when subjects are *overtly* primed with a construct, they tend to correct (or overcorrect) for its influence on their immediate subsequent judgements. She then observes that the mention of the possibility of error in the typical kinds of MEP-cases that epistemologists deal in cannot reasonably be described as subliminal. Indeed, it appears to be overt. But in that case we should not necessarily expect that mentioning error possibilities will lead to an increase in people's subjective probability judgements about their obtaining in the way that the availability heuristic explanation suggests, and indeed we may well expect the opposite effect. Call this the *spontaneous discounting objection*

§4.2.4 The cancelling objection

Nagel's forth objection is this: the error-salience phenomenon seems to be active in environments where we are reasoning carefully about epistemological issues (in the epistemology classroom, for example). So those who endorse the availability heuristic explanation will need to make the case that the heuristic is in operation in such environments. However, she notes, this suggestion is at odds with general facts about the workings of heuristics. Most heuristic tend to be diminished in effect in situation where we have "incentives for accuracy, or penalties for error, have been made self-conscious about our judgements, or expect to have to justify our judgements to others." (2010: 299) These conditions seem to be precisely though found in the epistemology classroom. In such contexts people tend to rely more on systematic processing and less on heuristics. So unless the heuristic that explains the error-salience phenomenon is especially robust or difficult to cancel, we should not expect it to have a substantial on our knowledge attributions and denials. And we may add (though Nagel does not) that the evidence suggests that the availability heuristic is not unusually robust. Call this the *cancelling objection*.

§4.2.5 The probability objection

Nagel's last objection has a more conceptual flavour. She objects is that it is unclear that a perceived non-zero probability of error makes the relevant kind of difference to knowledge attributions. People can, it seems, judge that the probability of the unrealized error possibility is very low, yet be inspired to deny that the subject knows. The persuasiveness of sceptical strategies, Nagel observes, seem to depend on the sceptic being able to persuade us that things would look the same to the subject in the world in which they falsely believe that p.

Persuasiveness doesn't seem to track probabilities. But in that case, it is unclear why the availability heuristic's leading people to overestimate the probability of the false belief possibility obtaining would result in them judging that the subject does not know. Furthermore, it is, Nagel argues, unclear why a slight increase in the apparent probability of the false belief possibility obtaining would be enough to sway people from judging that the subject knows to judging that he does not know. The problem is exacerbated by the fact that people will often judge that a subject lacks knowledge that p when the probability of p being false is explicitly stated as being very small - for example, in lottery cases - but continue to attribute knowledge of propositions that were formed using methods with a higher degree of fallibility. Given this, one might wonder why an increase in the perceived probability of epistemic danger would lead people to deny that a subject knows. Call this objection the *probability objection*.[42]

§4.3 How does the simulation heuristic explanation fare against Nagel's objections?

We have five objections to the availability heuristic explanation of the error-salience phenomenon from Nagel: the *imaginability* objection; the *comparative imaginability* objection; the *spontaneous discounting* objection; the *cancelling* objection; and the *probability* objection. This section addresses the question: how does the simulation heuristic explanation fare against these objections?

First though, I will make some remarks on the persuasiveness of the objections. I am persuaded by the imaginability objection, the spontaneous discounting objection, and the probability objection. Together they appear to me to torpedo the availability heuristic explanation, and I have nothing further to add to them. However, matters are more complex when it comes to the comparative imaginability objection, and the cancelling objection. With respect to the comparative imaginability objection, as Gerken (2013) points out, there seems

[42] Nagel offers a second conceptual objection that focuses specifically on the prospects of appealing to the availability heuristic explanation to defend invariantism, rather than simply to explain the error-salience phenomenon. Since the purpose of this essay isn't to defend invariantism, we will skip it.

to be something of a consensus amongst epistemologists that, *contra* Nagel, it is harder to get people to go negative with their knowledge attributions by mentioning far-fetched, recherché, and hard to imagine alternatives, than it is with commonplace, easier-to-imagine, alternatives (MacFarlane 2005, et al). Anyone who has taught an undergraduate epistemology class will know that it is a better idea to try to generate scepticism in one's students about whether S knows that her car is parked outside by pointing out that she can't rule the possibility that it has been stolen than by pointing out that she can't rule out the possibility that it has spontaneously reorganized into the form of a giant lizard. (The example is MacFarlane's). No empirical work has yet been done on the relationship between knowledge attribution and the comparative ease or difficulty of imagining mentioned alternatives. The matter should be studied. But we can, with at least a reasonable degree of confidence, proceed on the assumption that, *contra* Nagel, harder-to-imagine alternatives are less effective at generating negative knowledge attribution judgements than easier-to-imagine alternatives. For this reason, I do not find the comparative imaginability objection persuasive. And from here onwards I will not take it to be a desideratum on the simulation heuristic explanation that it is able to avoid Nagel's objection.

I will also not take it to be a desideratum on the simulation heuristic explanation that it is able to avoid the cancelling objection. The reason for this is that, whilst we should expect to see diminished reliance on non-robust heuristics in the circumstances that Nagel describes *when there are more methodical means by which to arrive at a judgement on the question at hand*, it is a striking fact about knowledge attribution that we *do not* have any methodical means by which to form judgements about whether a given subject knows that p. All we have to go on is our intuitive judgement. Accordingly, it is hard to see how whatever heuristics (if any) that have an influence on knowledge attribution would be diminished in effect, even if the assessment of whether a subject knows is being performed in circumstances under which reliance on non-robust heuristics is usually diminished. Thus, the cancelling objection misses the mark.

So how does the simulation heuristic explanation fare against the probability objection, the imaginability objection, and the spontaneous discounting objection? I see no difficulties for the simulation heuristic explanation analogous to those facing the availability heuristic explanation with respect to the probability objection. The simulation heuristic explanation makes no reference to probabilities, and makes no claim that people's judgements about the

probability of error possibilities explains the error-salience phenomenon. So there is no apparent reason why the fact that people can, at least ostensibly, continue to judge that the probability of the unrealized error possibility is very low and at the same time be moved to deny that the subject knows would create trouble for the simulation heuristic explanation. Nor, for the same reason, is there any apparent reason why trouble would be created by the fact that people will often judge that a subject lacks knowledge of the truth of a proposition when the probability of it being false is explicitly stated as being vanishingly small, but continue to attribute knowledge of propositions that were formed using methods with a higher degree of fallibility.

According to the imaginability objection, when events of a certain type are hard to imagine, the influence of the availability heuristic is to make people judge them to be *less*, rather than more, likely to obtain. And since it is far from clear that the kinds of MEP cases that philosophers and the folk have interacted with describe easy, rather than hard, to imagine, possibilities, it is correspondingly far from clear that the error-salience phenomenon can be explained by appeal to the workings of the availability heuristic. The objection clearly doesn't carry directly over to the simulation heuristic explanation. That explanation makes no reference to the likelihood of outcomes in its account, so the fact that people will judge hard-to-imagine events as less likely has no bearing on the plausibility of the explanation.

However, one might wonder: does a roughly analogous objection arise for the simulation heuristic explanation? Not obviously. Whilst there is good evidence that counterfactual possibilities that are easier to imagine tend to be mutated more than those that are hard to imagine (Byrne 2002, 2005), this fact doesn't obviously threaten the simulation heuristic explanation. The only inference we could reasonably make from the fact that hard to imagine possibilities are mutated less than easy to imagine possibilities is that there will be a greater tendency for people to deny knowledge in MEP cases where the error possibility is easy to imagine than there is in MEP cases where the error possibility is hard to imagine. But, as just discussed, it seems quite likely that this is what actually happens. Suppose that most of the MEP cases that epistemologists discuss do in fact describe error possibilities that are hard imagine. The only way that this would cause problems for the simulation heuristic explanation would be if, when people attended to hard to imagine possibilities, they became *even more confident* that things couldn't have been otherwise to how they were than their non-MEP counterparts who haven't had their attention drawn to *any* other possibilities. But there is

no evidence for or against this being the case. The matter will need to be studied further. But unless such a patterning is found, there will be no analogous imaginability objection against the simulation heuristic explanation.

What about the spontaneous discounting objection? Does the phenomenon of spontaneous discounting undermine the prospects of the simulation heuristic explanation? Nagel's concern was that when there is an obvious explanation for availability, people will tend to discount it in forming their judgements. And since MEP cases seem to be cases when the explanation for increased availability *is* obvious, it is unlikely that the error-salience phenomenon can be explained by the influence of the availability heuristic. One parallel concern for the simulation heuristic explanation would be if, when there is an obvious explanation for the increased salience of a counterfactual aspect of the situation, subjects will, as a result, tend *not* to mutate it to the same extent that they do when there is no obvious explanation for increased salience, *and* to a lesser extent than those who have not had their attention drawn to it The question of whether or not this is the case has not been studied. So more work will be needed here before we can draw any conclusions. (There is, it seems, no *a priori* reason to expect it to happen however). Another concern might be that when there is obvious explanation for the increased salience of an alternative, people will tend to discount the *relevance* of the salience. If so, then even if they come to recognise that the subject would falsely believe that p in this possible world, they will not be lead to deny that the subject knows that p because they will not treat this world as an epistemically relevant one. This possibility has not yet been studied. So more empirical work will need to be done here too.

§4.4. Nagel's hypothesis: Epistemic Egocentrism

It is not clear that any of Nagel's objections to the availability heuristic explanation cause trouble for the simulation heuristic explanation (though more empirical work will be required before we know where exactly it stands). However, having criticised the availability heuristic explanation, Nagel goes on to present her own hypothesis about the psychological mechanisms producing the error-salience phenomenon. This section presents Nagel's hypothesis. The next section discusses the relative strengths and weaknesses of my hypothesis and Nagel's, and presents outstanding empirical questions that will need to be studied in order

for us further our understanding of psychology behind the error-salience phenomenon.

Nagel's account appeals to a psychological bias known as *epistemic egocentrism*. Work by, amongst others, Royzman *et al.* (2003) and Birch *et al.* (2004), has shown that people are poor at representing the perspectives of those with less information than themselves. The result of this is that they often draw on privileged information when evaluating other's judgements. For example, Baron and Hershey (1988) tested subjects by having them evaluate fictional medical and financial decisions. Baron and Hershey provided the subjects with the information on which decisions were supposedly made, and also on the supposed outcomes of the decisions. Baron and Hershey found that although most subjects believed that the information about the supposed outcomes had no effect on their evaluation of the decision, in reality their evaluations were strongly biased by the outcome given. Despite recognising that the people whose decisions they were evaluating did not share their knowledge of the outcomes when they made their decisions, when they were asked about the quality of the decisions, the subjects represented the decision-makers egocentrically, as though they did have this knowledge.

Nagel's proposal is that the bias of epistemic egocentricity might explain the error-salience phenomenon as follows:

> "If we naturally evaluate the statements and judgements of more naive subjects as though they shared our privileged beliefs and concerns, without being aware that we are doing so, we may unfairly penalize them for failing to respond to those concerns as we would, just as Baron and Hershey's subjects penalized the fictitious doctors whose patients died from a generally very beneficial procedure with a small risk of death" (2010: 303)

How would this apply to the (a) and (b) cases? Nagel suggests that, since in (b) but not (a), concerns about the possibility of tricky lighting have been raised, readers will mistakenly evaluate John as if he shared those concerns in (b). Once they have done so, they will interpret the fact that he did not check the colour of the lighting, despite having a concern as to its colour, as a sign of compromised or motivated belief formation. The typical hallmarks of compromised or motivated belief formation - haste, distraction, wishful thinking etc. - are conditions which typically lower the accuracy of judgement. When the accuracy of one's judgement is compromised, one seems to be a mere believer, rather than a knower. Thus

readers of (b) will display a greater tendency than readers of (a) to deny that John knows, exactly in line with the error-salience phenomenon. The same will go for all MEP-cases.

§4.5 Comparing the simulation heuristic explanation and the epistemic egocentrism explanation

This section examines the respective merits of the epistemic egocentricity explanation and the simulation heuristic explanation. An important first observation is that, just as before with the availability heuristic explanation, the simulation heuristic explanation and the epistemic egocentricity explanation need not be seen as rivals, for it is entirely possible that they *both* contribute to the error-salience phenomenon (Nagel makes no claim that her explanation is exhaustive).

What can be said in favour of the epistemic egocentricity explanation? First off, it appears to avoid most of the objections facing the availability heuristic explanation. If the error-salience phenomenon is produced not by increased availability, but rather by the mistaken projection of concerns on to subjects, then the fact that when events of a certain type are hard to imagine, people tend judge them to be *less*, rather than more, likely to obtain, would appear to be irrelevant to whether they will display more of a tendency to deny knowledge in MEP-cases than in non MEP-cases. Additionally, as Nagel points out, unlike the availability heuristic, the epistemic egocentricity bias is known not to be subject to spontaneous discounting (Pohl & Hell 1996, Krueger & Clement 1994). So the epistemic egocentricity explanation is not vulnerable to the spontaneous discounting objection. Nor is there any reason to think that it will be vulnerable to the probability objection - concerns aren't a function of probabilities.

Moreover, the epistemic egocentricity explanation has been subjected to empirical testing by Alexander, Gonnerman, & Waterman (forthcoming) (hereafter 'AGW'). AGW observed that the epistemic egocentricity explanation predicts that there should be relatively little difference between *narrator cases* - cases where the relevant information about the possibility of error is shared only with the reader, and not the subject - and *subject cases* - cases where the possibility of error is shared with the subject of the case (for example, by including another party in the story who tells the subject that, say, the table would have looked red to him if he

had been looking at a white table under a red light). AGW tested this prediction by comparing the strength of agreement in readers with the statement 'John knows that the table is red'. across narrator cases and subject cases. In line with the prediction of the epistemic egocentricity explanation, they found no statistically significant difference in readers mean strength of agreement with the statement. This result is compatible with the epistemic egocentricity explanation.

AGW also hypothesised that if epistemic egocentricity is producing the error-salience phenomenon, then we should find a negative correlation between the projection of our concerns onto others and our willingness to attribute knowledge to them. They tested this hypothesis by presenting participants with Nagel's (b) MEP-case, and asking them to indicate their strength of agreement with (1) the claim that John knows that the table is red, and (2) the claim that John is considering the possibility that he is looking at a white table under a red spotlight. They found a modest negative correlation. It is, they note, unclear what this tells us about the epistemic egocentricity explanation. One possibility, they suggest, is that it is exactly what we should expect on that explanation, since the standard line on epistemic egocentrism is that people are unaware of the fact that it has occurred (Baron & Hershey 1988, Camerer *et al.* 1989). If people are unaware that they are projecting concerns, then we shouldn't expect to find a strong negative correlation simply by *asking* subjects whether they thinking that John shares the concern, since they will presumably have limited access to their projections. If this is right, AGW note, then if we are to find good evidence for the epistemic egocentricity explanation, we need to find better ways of measuring the largely unconscious projection. However, they also consider some other possibilities. Perhaps for example, people are aware of their projections, but raise the standards for knowledge attribution in conversational contexts that mentioned unrealized error possibilities in non-uniform ways. More work is needed here.

The epistemic egocentricity explanation certainly has things to be said in favour of it. However, two concerns remain outstanding. Firstly, as discussed in §7, there is some reason to think that, *contra* Nagel, it is harder to elicit knowledge denials in MEP-cases where the mentioned unrealized error possibility is hard to imagine, than it is in MEP-cases where the mentioned unrealized error possibility is easier to imagine. Epistemologists do a better job of making sceptics of their students by pointing out that you can't rule out that your car was stolen than by pointing out that it could have spontaneously reformed in the form of a giant

lizard. If this is right, can the epistemic egocentricity explanation account for it? The matter is unclear. There is no obvious reason to expect that the fact that the counterfactual possibility one is concerned about is hard to imagine would lead one to be less inclined to deny that a subject who shares the concern, but does nothing to rule it out, knows. But in that case, there is no obvious reason to expect that the epistemic egocentricity explanation does predict the (apparent) fact that harder-to-imagine cases elicit less of a tendency towards knowledge denial. And that might be seen as a mark against it. However, caution is needed here. Firstly, as noted earlier, no empirical work has yet been done to show that harder-to-imagine MEP-cases result in less knowledge-denial than easier-to-imagine MEP-cases. Secondly, ease of imagination effects can have subtle and surprising interactions with other effects.[43] So it would be hasty to draw a priori conclusions. The matter should be studied further.

The second worry is this: the epistemic egocentricity explanation relies on the assumption that in MEP cases like case (b) merely *mentioning* the possibility of error is sufficient to raise it as a concern in the reader's mind. But it is not obvious that this is the case. Mentioning something isn't the *same thing* as expressing a concern about it, so if the epistemic egocentricity explanation is to be persuasive we will have to be given good reason to think that in MEP cases mentioning tends to trigger concern. Again, more work is needed here.

The simulation heuristic explanation doesn't face these two worries. Since it makes no reference to concerns, it does not have to make good on the claim that mentioning triggers concern. And it seems quite clear that we *should* expect to find the phenomenon of harder-to-imagine cases eliciting less of a tendency towards knowledge denial if the simulation heuristic explanation is correct. It is widely recognised that people tend to mutate alternatives that are easy to imagine to a greater degree than those that are hard to imagine (Byrne 2002, 2005). So if the error-salience phenomenon is the result of people mutating the mentioned error possibility in MEP-cases and thereby coming to deny that the subject knows, then we should expect to see this tendency weakened when the mentioned error possibility is harder-to-imagine as compared to when it is easier-to-imagine. This is a point in favour of the simulation heuristic explanation.

What about AGW's studies? Do they tell against the simulation heuristic explanation? They do not. AGW found no difference in people’s tendency to attribute knowledge between narrator cases and subject cases. This result is in line with the epistemic egocentricity

[43] Thanks to Jennifer Nagel here.

explanation, but it equally well comports with the simulation heuristic explanation. In both the narrator and the subject cases, the reader's attention was drawn to the unrealized possibility that John would have believed that he was looking at a red table if he was looking at a white table under a red light. The cases are the same in this respect. And according to the simulation heuristic explanation, it is this fact about attention that causes people to deny that he knows. So on the simulation heuristic explanation, we wouldn't expect to see any difference in knowledge attribution between the narrator and subject cases. Thus AGW's first study equally well supports the simulation heuristic explanation as the epistemic egocentricity explanation.

What about AGW's second study? The simulation heuristic explanation does *not* predict that there will be a negative correlation between experimental participants projection of concerns onto others and their willingness to attribute knowledge. So *prima facie* the fact that AGW found a negative correlation would seem to be a blow to the simulation heuristic explanation. However, matters are not so straightforward. As AGW note, the negative correlation they found was very modest. AGW suggest that this might be what we should expect on the epistemic egocentricity explanation, because the standard line on epistemic egocentrism is that people are unaware of the fact that it has occurred. But if the influence of epistemic egocentrism really is opaque to those it affects, then it is unclear that we should expect to find *any* negative correlation of the sort that AGW found on the epistemic egocentricity explanation. If people don't realise that they are illegitimately projecting concerns on to the subject, then *asking them* whether they think the subject has the relevant concerns ought not to be an effective experimental method. So it is quite unclear what AGW's second study tells us, and what is producing the modest negative correlation. Insofar as these matters are unclear, we should be hesitant to take the results to undermine the simulation heuristic explanation.

In the next section I show how the simulation heuristic explanation integrates with Gerken's epistemic focal bias theory. However, before we proceed, it will be useful to sum up some of the outstanding empirical questions.

Before we can be confident that the error-salience phenomenon is the result of the influence of the simulation heuristic, we will need to know more about the nature of the simulation heuristic. Two point of interest have been identified: firstly, it should be investigated whether the simulation heuristic is subject to spontaneous discounting in such a way that this would

undermine the simulation heuristic explanation. Secondly, the matter of just what effect drawing people's attention to hard-to-imagine counterfactual aspects of a situation has on their counterfactual reasoning should be investigated.

One empirical question of interest to both the simulation heuristic explanation and the epistemic egocentricity explanation is the contrasting effects of presenting people with easy-to-imagine MEP-cases and hard-to-imagine MEP-cases. Are people more willing to attribute knowledge when the mentioned unrealized alternative is hard to imagine than when it is easy to imagine? Epistemologists have tended to assume that they are. But the matter can, and should, be tested empirically. If the assumption turns out to be correct, then it will be important for the epistemic egocentricity explanation to investigate the effects of difficulty of imaginability on epistemic egocentricity. If the assumption turns out to be incorrect, this will be a strike against the simulation heuristic explanation.

§4.6 Gerken on epistemic focal bias

Mikkel Gerken (2013) has also offered a proposal about the psychological processes generating the error-salience phenomenon. In this section I briefly present Gerken's proposal, and show how it does not compete with, but rather complements, the simulation heuristic explanation. Gerken's account, which he calls 'epistemic focal bias', derives from the following two principles:

> ***Principle of contextual salience***: Normally, for an agent, A, q is a contextually salient alternative to S's knowledge that p iff A processes q as an epistemically relevant alternative to S's knowledge that p
>
> ***Principle of epistemic satisficing:*** Normally, an agent A, forms epistemic judgements on the basis of a prima facie reason that is arrived at by processing only a limited part of the evidence that is available to A.

We discussed the principle of contextual salience in §3. As Gerken says:

> "Often the psychological salience of some feature of a scenario affects our cognitive

> processing whether or not that feature is relevant to the task at hand. This psychological feature is reasonable since, for many tasks, the contextual salience of something can amount to defeasible evidence for its relevance to the task (Bach 1985, 2010: §5; Gigerenzer and Todd 1999; Stanovich 2009). If so, processing salient features of a scenario as epistemically relevant is normally a cognitive practice that exemplifies bounded rationality since this mode of processing makes good use of the subject's limited cognitive resources (Stein 1996)." Gerken (2013: 50)

Gerken's reason for postulating the principle of epistemic satisficing is that it is a very general fact about human cognition that people will often reach a judgement via a process that comes to a halt once a *prima facie* reason has been found (Simon 1955, 1983).

Gerken observes that if the principle of contextual salience and the principle of epistemic satisficing both hold, then we should expect the error-salience phenomenon. When an alternative, q, has been made salient, as happens in MEP-cases (but not in non MEP-cases), according to the principle of contextual salience people will normally regard this as incompatible with S's knowing that p, unless they regard S as being able to rule out q. So in MEP-cases people will normally acquire a *prima facie* reason to regard S as a non-knower. They will not acquire any such reason in non MEP-cases. According to the principle of epistemic satisficing, the processing normally comes to a halt once a *prima facie* reason for making a judgement has been arrived at. That is, people will normally not engage in further processing by critically assessing the salient alternative in the light of her background beliefs or further evidence. Rather, they will normally judge that S does not know in MEP cases. In this way, the error-salience phenomenon will be produced.

The epistemic focal bias account is perfectly compatible with the simulation heuristic explanation. Indeed the two complement one another.[44] The simulation heuristic explanation primarily offers an account of *why people will display more of a tendency to recognise that there is a possible world in which the subject falsely believes that p in MEP-cases than in non MEP-cases*. The epistemic focal bias account, on the other hand, offers an explanation for *why this disparity in recognition will lead people to deny knowledge in MEP-cases but not in non MEP-cases*. Since the two accounts focus on different parts of the set of psychological processes that produce the error-salience phenomenon, they do not compete with one another.

[44] Gerken notes that the epistemic focal bias account and the epistemic egocentricity explanation may also complement one another (2013: 59 f26)

And in fact they complement one another, since combined they offer us a more thorough explanation of the phenomenon than either does alone. For example, supposing that it is harder to get people to go negative in their knowledge judgements by mentioning harder-to-imagine error possibilities than it is with easier-to-imagine error possibilities, the epistemic focal bias account alone lacks the resources to explain why this is. The simulation heuristic explanation, by contrast, predicts it.

V. The Semantics Of 'Knows' Again

As discussed in §2, Williamson, Hawthorne, Nagel, and Gerken, all take their proposals to lend some support for invariantism about the semantics of 'knows'. In each case, the argument takes the same form: the existence of the error-salience phenomenon appears to pose a challenge to invariantism, since if invariantism were correct, it would show that people commit systematic errors in their knowledge attribution. However, if it can be shown that we should expect there to be such errors *even if* 'knows' has an invariantist semantics, because they are the result of well known psychological heuristics or biases, then the case against invariantism is at least partially undermined. And of course this is what the invariantists try to show - Williamson and Hawthorne with the availability heuristic explanation, Nagel with the epistemic egocentricity explanation, and Gerken with the epistemic focal bias account.

One might also think that the simulation heuristic explanation similarly blocks the argument against invariantism from the error-salience phenomenon. The thought would be that people rely on mental simulation to form judgements about whether the subject's belief is sufficiently modally robust to count as an item of knowledge, but because mental simulation is influenced by a heuristic which leads people to mutate salient aspects of a scenario more than non-salient

aspects, people are systematically led astray in their knowledge attributions when presented with salient, but epistemically irrelevant, error possibilities.

I have reservations about appealing to the psychological processes producing the error-salience phenomenon to defend invariantism. Though I certainly don't have a knock-down argument against the prospects of doing so, I will in closing, briefly sketch a line of thought according to which we should expect the error-salience phenomenon producing psychological processes to be constitutive of the concept of knowledge. We begin with the observation that according to traditional Bayesian decision theory, rational decision makers maximize expected utility, where expected utility is a function of the value of various possible outcomes, and the probability of their occurring. However, as Ross and Schroeder (2014) point out, for almost every proposition one should, and typically will, have a non-zero credence that it is true, since there is nothing that we can be completely certain is false. If we were to factor all of these non-zero probabilities into our decision making, then coming to a decision about how to act would very often be an unmanageably complex task for creatures, like ourselves, who have limited cognitive capacities and often need to engage in decision making in time-pressured circumstances. For this reason it has long been recognised by decision theorists - Savage (1972) and Joyce (1999) - and more recently by epistemologists - Weatherson (2005) and Ross and Schroeder (2014) - that when we are engaged in practical reasoning we must take some uncertain propositions to be certain for the purposes of decision making in order for decision making tasks to be manageable. One recent proposal is that our system of practical reasoning is governed by a norm stating that one should treat p as certain for the purposes of practical reasoning only if one knows that p. This is, of course, controversial.[45] Suppose, however, that it is correct. In that case, it seems that one central role for the concept of knowledge is to simplify decision making tasks by allowing us to take as certain in our practical reasoning uncertain propositions that are known to be true.[46] This is where the concern arises for invariantism. Given that we face a huge number of decision making tasks every day, we need a way of making good decisions in a cognitively cheap manner. If the function of the concept of knowledge is to help us to do this, by enabling us to simplify decision making tasks, then we will need a concept of knowledge that allows us to make accurate knowledge attributions in a cognitively cheap manner, if we are to be rational

[45] Brown (2008), Neta (2009), Reed (2010), Schiffer (2007), and Gerken (2011), amongst others, dispute the claim. This issue is discussed in greater detail in the essays 'Excuses and Epistemic Norms', and 'From Moore's Paradox To The Knowledge Norm Of Belief And Beyond' in this dissertation.
[46] See the essay 'From Moore's Paradox To The Knowledge Norm Of Belief And Beyond' in this dissertation.

decision makers. And there is some tension between this need, and the invariantist suggestion that reliance on heuristics leads to systematic error in our knowledge attribution. If heuristic psychological processes have a tendency to lead us astray in our knowledge attributions in the way that the discussed proposals suggest when coupled with invariantism, then accurately judging whether a subject - either oneself of another - knows or not, will be a fairly hard cognitive task, requiring correction from system 2 psychological processes, rather than sole reliance on system 1 processes. Thus, making accurate knowledge attributions will be a relatively laborious and cognitively expensive task if invariantism is correct as compared to if contextualism is correct, for on contextualism, relying on such heuristics will tend *not* to lead one astray in one's knowledge attributions. For this reason I think it is unclear that an invariantist concept of knowledge would be fit to play one of the central normative roles that it is designed to play. The tension is between the fact that we employ the concept of knowledge to ease the cognitive burden of decision making and the fact that on invariantism accurate knowledge attribution will be a cognitively burdensome task. The tension appears to be lessened if our concept of knowledge is contextualist, and this, one might think, is a *prima facie* reason to prefer contextualism if, as the proposals discussed here suggest, knowledge attribution is influenced by heuristic processes.

As I said, this is hardly a conclusive argument. It is possible that even if invariantism is correct, sufficiently accurate knowledge attribution is cognitively cheap *enough* for the concept of knowledge to play the required role. Nevertheless, there does appear to be some cause for concern. This issue should be explored more, but this task is beyond the scope of this essay.

VI. Conclusion

Questions about what psychological processes produce the error-salience phenomenon are complex and relatively under-explored. Much more work will need to be done before we can

confidently draw any conclusions. Nevertheless, there is, I have argued, some reason to think that the phenomenon might be at least in part explained as the result of the influence of the simulation heuristic on knowledge attribution. The goal of this essay has been to sketch how such an explanation would go. This lays a foundation for further investigation.

4. Excuses And Epistemic Norms

§I. Norms, Justifications, and Excuses

§1.1 The Knowledge Norm of Practical Reasoning

Recently it has been popular to argue that there are epistemic constraints on appropriate practical reasoning. That is, in order for a chain of practical reasoning to be permissible, from an epistemic point of view, the reasoner must bear a certain epistemic relation to the propositions employed in the reasoning. Just what this epistemic relation might be is a matter of controversy. John Hawthorne and Jason Stanley (2008) defend the following principle:

> ***RPK:*** Where one's choice is p-dependent, it is appropriate to treat the proposition p as a reason for acting iff you know that p

RPK states that, when p is relevant to practical decision making, knowing that p is necessary and sufficient for appropriately treating p as a reason for acting, where appropriateness should be understood in terms permissibility: it is epistemically permissible to treat p as a reason for acting iff you know that p.[47] The necessity direction claim of *RPK*, which will be the focus of this essay (call it the 'K-Norm′), has come in for criticism.[48] A number of authors have offered apparent counterexamples to the K-Norm and argued that, as a result, we should prefer weaker principles requiring subjects to merely stand in some kind of justification or justified belief relation to p in order to appropriately employ it in practical reasoning (call these 'J-Norms'). The purpose of this essay is threefold. Some have objected to appeals to the justification/excuse distinction made by K-Norm proponents in response to the would-be counterexamples. In the first section of the essay I reply to these objections, arguing that they

[47]Elsewhere Hawthorne and Stanley offer up variations of this idea: 'One should act on p only if one knows that p' (Stanley, 2005: 9); 'If p is practically relevant, it is acceptable to use the premise that p in one's practical reasoning only if one knows that p, and (at least in many cases) unacceptable to use the premise that p in one's practical reasoning if one doesn't know it' (Hawthorne, 2004: 30). Fantl and McGrath (2009) also defend several knowledge-action principles. The most recent and important of these is 'KJ': 'If you know that p, then p is warranted enough to justify you in Φ'ing, for any Φ'. In this essay I will focus only on the formulation found in Hawthorne and Stanley (2008).

[48] As has the sufficiency direction. See Brown (2008)

are, variously, unpersuasive and inconclusive. In the second section of the essay I explore the question of whether rivals to the K-Norm proposed by these critics are subject to would-be counterexamples analogous to those facing the K-Norm, and answer in the affirmative. I argue that, as a consequence of this, the rival proposals will likewise require an appeal to the justification/excuse distinction. Finally, in the third section, I argue that, whilst this fact undermines the force of the J-Normers objections to the K-Norm, we have nevertheless as yet been given no reason at present to prefer the K-Norm to the rival norms, and vice versa, since we have been given no reason to prefer to draw the line between justification and excuse where K-Normers would have us draw it rather than where J-Normers would, and vice versa.

§1.2. The K-Norm and Alleged Counterexamples

Hawthorne and Stanley (2008) offer two kinds of considerations in defence of the K-Norm. Firstly, they observe that there is a close connection between the concept of knowledge and out everyday assessments of the behaviour of others. For example if Hannah and Sarah are trying to find the restaurant at which they have reservations, if instead of asking for directions, Hannah goes on her hunch that the restaurant is down a street on the left, and turns out to be mistaken, a natural way for Sarah to criticise Hannah's behaviour is by saying 'You shouldn't have gone down this street; you didn't know that the restaurant was here'. The K-Norm, they observe, explains the appropriateness of Sarah's criticism. If it is permissible to act only on what you know, then the reasoning that led Hannah to act as she did was impermissible and she is, *ceteris paribus*, thereby criticisable on these grounds. Similarly, if John decides not to buy health insurance anymore, reasoning that he won't need it because he won't get ill, his mother may appropriately berate him for his decision on the grounds that he doesn't know that he won't fall ill. Again, the K-Norm explains the appropriateness of the criticism, in the same way that it does for Hannah. Moreover, Hawthorne and Stanley point out, knowledge also interacts with conditional orders in a way that supports the K-Norm. "Suppose a prison guard is ordered to shoot a prisoner if and only if they are trying to escape. If the guard knows someone is trying to escape and yet does not shoot he will be held accountable. Suppose meanwhile he does not know that someone is trying to escape but shoots then anyway, acting on a belief grounded in a baseless hunch that they were trying to

escape. Here again the person will be faulted." (2008: 572) Our common practice, Hawthorne and Stanley suggest, is to require knowledge of the antecedent of a conditional order in order to discharge it. Again, they observe, the suggestion that the K-Norm governs practical reasoning explains this fact. Lastly, they argue, blame and judgements of negligence interact with knowledge in a way that provides support for the K-Norm. For example, if a doctor uses a needle that he does not know to be clean, then he is *prima facie* negligent.

The second consideration that Hawthorne and Stanley offer in defence of the K-Norm stems from the following scenario. Suppose that you are offered 1 cent for a lottery ticket that cost $1, in a 10,000 ticket lottery with a $5,000 first prize, and you reason as follows:

(1) I will lose the lottery
(2) So, if I keep the ticket I will get nothing
(3) If I sell the ticket I will get 1 cent
(4) So, I ought to sell the ticket

This piece of practical reasoning is, Hawthorne and Stanley note, obviously defective. If one has only probabilistic evidence for the claim that one's ticket will lose, then it is clearly inappropriate to treat the proposition that the ticket will lose as a premise in one's practical reasoning, even if the proposition is in fact true, and enjoys strong probabilistic support. The most natural explanation of this, they suggest, is because, as the K-Norm claims, it is impermissible to treat p as a reason for acting unless one knows that p, and one cannot know merely on the basis of the probabilities involved that one's lottery ticket has lost.

There is a great deal that can (and has) been said about these arguments for the K-Norm. For the purposes of this essay though, we can restrict our attention to one kind of objection to the view; namely that the K-Norm is subject to counterexamples from cases where subjects fail to know the relevant proposition, but it nevertheless seems inappropriate to criticise them for treating it as a premise in practical reasoning. Such cases come in two forms: Gettier cases, and justified false belief cases. Jessica Brown (2008) offers a good example of the first:

> ***TIMETABLE:*** *Suppose...that S leaves the office at 12.00pm in order to meet her partner for lunch at 1pm. S believes truly that there is an express train at 12.20pm which would allow her to arrive in time to make lunch. Further, this belief is justified:*

> *S checked the train timetable on the internet just before leaving the office. In fact, unbeknownst to S, she is in a Gettier situation:a hacker has got into the train website and for a joke has replaced all of the current timetables with last season's timetables. Luckily for S, according to both the old and new timetables, there is an express at 12.20pm. So, S's belief is a case of true justified belief but not knowledge. (2008: 172)*

Suppose that S acts according to her belief that there is a 12.20 express train. Brown asks us to consider whether it is appropriate for her to do so. By Brown's judgement, it is. Intuitively, it doesn't seem appropriate for S's partner to criticise her actions, or her relying on her belief that there was a 12.20 express train. After all, S wasn't to know that the website had been hacked. It would be *unreasonable* for S's partner to criticise her by saying, for example, 'You shouldn't have left so late, as you didn't know there was an express at 12.20'. Brown takes the fact that such criticism would be unreasonable to show that, *contra* Hawthorne and Stanley, knowledge is not the norm of practical reasoning. As a general point, Brown argues, cases like *TIMETABLE* show that when a subject fails to know that p because her true belief is Gettiered, but would know that p were she in a non-Gettier case, the fact that her belief is Gettiered does not change the appropriateness of her relying on p in her practical reasoning. If Brown is right about this, then the K-Norm is false. Ram Neta (2009) makes the same point, as does Mikkel Gerken (2011).

Gettier cases are not the only way that the K-Norm has been called into question. As Lackey (2007), Neta (2009), and Hawthorne and Stanley (2008) recognize, there are cases where a subject has a great deal of justification for p, and hence every reason to think that they know that p, but the subject doesn't know that p because p is false, yet it seems as though it would be wrong to criticise the subject for their behaviour when they act on p. An adaptation of the above case will serve to illustrate the point. Suppose the case is exactly the same, except that there is no express train at 12.20 - the old timetable is not the same as the new timetable when it comes to 12.20 trains - and S misses her lunch. Were S's partner to complain about her behaviour, she would, it seems, be well within her rights to respond that she had every reason to think that there was a 12.20 express - after all, the website said there would be, and how was she to know it had been hacked? - and that this legitimised her behaviour. Arguably, in this case, despite S not knowing that p and p being false, it was nevertheless permissible for S to rely on p in her practical reasoning. If it was, then the K-Norm is false, since there are non-knowledge conditions under which it is epistemically permissible to employ p in one's

practical reasoning.

§1.3. Responding To the Counterexamples: The Justification/Excuse Distinction

We have, then, two types of apparent counterexamples to the K-Norm. Proponents of the K-Norm are well aware of these. In an attempt to vitiate their force, Hawthorne and Stanley follow Williamson (2000) in appealing to the distinction between *justifiably* Φ-ing and *excusably* Φ-ing. They argue that Gettier cases and justified false belief cases are cases where the subject does not appropriately employ p in their practical reasoning - in line with the verdict of the K-Norm - but is to be *excused* for failing to meet the norm's conditions, because they have every reason to believe that they do meet the norm's conditions. Since their behaviour is excusable, they are not to be blamed or criticised. But, Hawthorne and Stanley argue, merely excusable behaviour should not be confused with justified behaviour, and the intuition that Gettiered subjects and justified false belief subjects are not open to criticism for acting as they do should, *contra* the critics, not be taken to show that they are fully justified in acting as they do. Moreover, Hawthorne and Stanley think, not only does the appeal to the distinction vitiate the force of the Gettier and justified false belief would-be counterexamples to the K-Norm, but also that the distinction is unavoidable. On this point they remark:

> "...as Tim Williamson has emphasized, [the need to distinguish between justified Φ-ing and excusable Φ-ing] will arise whatever one's normative theory, given that no conditions are luminous...In general, luminosity failure makes for confusion or at least hesitancy in our normative theorizing...After all...excusable lapses from a norm are no counterexample to that norm" (2008: 578)

The idea here is that, since no conditions are luminous - where a condition C is luminous iff for every case A, if in A condition C obtains, then in A one is in a position to know that C obtains - it will always be possible for a subject to justifiably, but falsely, believe that they satisfy the conditions specified in a norm. Since the subject has 'done their best', they are excused from transgressing the norm. But they nevertheless fail to satisfy it. In that case, one

might think, there are no benefits to be had from moving to weaker justification norms, and the would-be counterexamples to the K-Norm fail to imperil the theory. We will turn to assessment of these claims about the universality of a need to appeal to the justification/excuse distinction shortly. Before we do, it should be noted that a number of critics of the K-Norm have found the justification/excuse response unpersuasive, and used their objections, along with the would-be counterexamples, as a springboard to motivate rival proposals.

§1.4 Objections to Appealing to the Justification/Excuse Distinction & Replies

Jon Kvanvig's (2010) objection to Hawthorne and Stanley's reply can be found in the following passage:

> "...the world is full of rules and requirements, 'musts' and 'oughts', but what we have a need for are the epistemic underpinnings relevant to the fundamental questions of what to do and what to believe. If we wish, we can put the answers to such questions in terms of normative terminology: things we should do or should believe, things that are the right thing to do or the right thing to believe. If we do so, we should understand the relevant normativity in terms of the basic questions regarding what to do and what to believe...[one problematic possibility] is to offer multiplicity where we seek simplicity. Instead of offering one theory that tells us what to do and what to think, this approach generates multiple answers, so that doing A or thinking B can be prescribed in terms of one normative notion, but proscribed in terms of another. In the face of such complexity, a Socratic response is appropriate: we want to know one thing - what to do and what to think - and we get multiplicity. No, we want just one thing. It is really important that we get it, and we will not pay much attention to theories that do not answer this fundamental concern. Moreover, multiple answer theories are simply unhelpful in this context, leaving only perplexity in their trail." (2010: 237-238)

Kvanvig takes the K-Norm, shored up with an appeal to the justification/excuse distinction, to

be a 'multiple answer theory', and hence undesirable. It is undesirable because appealing to the distinction conflicts with the fundamental goal of normative theorizing: to answer the central questions of the egocentric predicament - in this case the question being: 'when should I employ a proposition as a premise in practical reasoning?'. Should this objection make us hesitant to accept the K-Norm, given that it must appeal to the justification/excuse distinction in order to explain our judgements about Gettier and justified false belief cases? I'm sceptical. Taken on its face the objection belies a misunderstanding of the nature of the distinction proponents of the K-Norm wish to appeal to. Recall, the thought is that when one acts in accordance with the norm one's behaviour is justified, and when one blamelessly fails to act in accordance with it, one's behaviour is excusable. Arguably it is a mistake on Kvanvig's part to interpret this as a multiple answer theory. His complaint appears to be that a theorist touting such a theory fails to properly tell you what to do, since his answer, of the form 'You can do A, and you'll be justified, or you can do ~A, and you'll be excused', licenses conflicting options, and hence fails to give guidance on how to proceed. Interpreted in this way, Kvanvig's objection misses the mark. Such a theorist *has* told you what to do: A. Saying that one will be excused if one performs some act Φ is not, in any sense, to recommend, instruct, or advise that one Φ. Rather, to say that one will be excused if one Φ's is to inform that, under certain circumstances, one will not bear the adverse normative consequences that normally accompany Φ-ing. This is simply not the same thing as instructing, advising, or recommending Φ-ing. Excusable actions are wrong actions. They should not be done. Accordingly, one might think, it's hard to make sense of Kvanvig's complaint here, since there is no sense in which the proponent of the K-Norm is offering multiple answers to the question of what to do.

However, there may be a more serious concern in the vicinity. Arguably a subject, like S in *TIMETABLE,* who has every reason to think that they know that p, and is in a situation where p is relevant to their decision making, would be irrational *not* to treat p as a reason for acting. Suppose that S decides, on nothing more than a baseless hunch that it might be wrong, not to rely on the timetable in front of her but instead to phone the train company before deciding what to do. It is quite natural to think that in behaving as she does, S behaves irrationally in virtue of being overly cautious. In that case, given the plausible assumption that rationality has normative force we do indeed have a multiple answer theory, for on the one hand we have the K-Norm telling S that she shouldn't employ p as a premise in her reasoning, and on the other, rationality telling her that she should. What then is she supposed to do? Which

suggestion should win out? This issue is one that K-Normers will have to face up to. However we shouldn't simply *assume*, as Kvanvig appears to, that a satisfactory non multiple-answer theory is available. No doubt, were it possible to develop one, it would be theoretically desirable to do so. But as Hawthorne and Srinivasan (2013) have emphasised, and as we will see in further detail later, given the extent and degree of luminosity failure, it may be that we are simply *forced* into accepting a multiple answer theory, whether we like it or not. If so, then Kvanvig's objection will be impotent. As it stands, it is inconclusive.

Mikkel Gerken (2011) objects in a different way to Hawthorne and Stanley's appeal to the justification/excuse distinction in defence of the K-Norm. His complaint is that there is no workable fleshing out of its details. Quite reasonably, he demands that proponents of the K-Norm tell us just what the conditions are for one to excusably violate the norm. Running through several possibilities, he finds each wanting, and goes on to argue that the only plausible proposal ultimately collapses into his own *Warrant Account*, which states that:

> ***WA**:* In the deliberative context, DC, S meets the epistemic conditions on rational use of (her belief that) p as a premise in practical reasoning or of (her belief that) p as a reason for acting only if S is warranted in believing that p to a degree that is adequate relative to DC [49]

WA takes a deliberative context as, not the *de facto* circumstances S is in, but rather the circumstances S rationally presupposes herself to be in, where the notion of rationality in play is externally constrained. Gerken means 'warranted in believing' to be roughly synonymous with 'justified in believing'. So *WA* states that a minimal epistemic constraint on appropriate employment of p in practical reasoning is that the subject employing p is justified in believing that p to a degree adequate to the circumstances the subject rationally presupposes herself to be in. Whilst he acknowledges that in some deliberative contexts knowledge will be required, Gerken also thinks that in others mere justified belief will be sufficient.

Why does Gerken think that the only plausible account of the justification/excuse distinction in this context collapses into the *WA?* First he considers the following notion of excusability: *one is excused for failing to satisfy the K-Norm when one has every reason to think that one*

[49]Gerken actually expresses *WA* as a biconditional. Again, since our interest in this essay is only in necessity direction principles, I have amended the principle accordingly.

knows that p. He finds fault with this proposal, pointing out that very young humans most likely cannot entertain higher-order thoughts such as 'I know that p', and hence can't satisfy the excusability condition. Since, Gerken thinks, such subjects are capable of excusably failing to satisfy the epistemic norm of practical reasoning, this notion of excusability is flawed. The second possibility that Gerken considers is that *a subject is excused when they possess knowledge level propositional justification for p*.[50] He takes this to be even less appealing than the first proposal, since it is not clear how the presence of such justification could be *sufficient to excuse* given that a subject can possess propositional justification for p even when this justification has no effect on the reasoning that leads to their acting, and hence has no place in an explanation of why the subject acted as they did. Finally, Gerken considers a proposal he finds more promising: *a subject is excused when they justifiably believe that p*. Here he adds a basing requirement to the propositional justification requirement - the subject must possess doxastic justification for p. However, Gerken maintains that this approach simply collapses into his own WA, because:

> "It is compatible with a warrant approach to hold that the degree of warrant normally required for knowledge is what is normally required for action/practical reasoning. But a central reason for distinguishing between "warrant normally required for knowledge" and "knowledge itself" is this: In certain epistemically abnormal cases, the degree of warrant that would have met the warrant condition on knowledge in normal circumstances is insufficient for knowledge. In some such cases, the warrant may nevertheless be sufficient for rational action/practical reasoning. As mentioned, cases of warranted false belief and Gettier-style cases are paradigms of such epistemically abnormal cases. So, to say that knowledge is normally required for action/practical reasoning but that warranted belief will do in abnormal cases amounts to accepting a version of a warrant account compatible with WA" (2011: 542)

This line of thought is mistaken. Gerken assumes that proponents of the K-Norm will accept that the 'abnormal' cases are ones where the subject is *justified* in employing p in their practical reasoning - in his words, warranted belief "will do" in abnormal cases. If K-Norm proponents did accept such a thing, then Gerken would be right that the K-Norm collapses into his own *WA*, since they will deliver the same verdicts in the same cases. However, as we have already seen, proponents of the K-Norm will not accept this. They claim - unlike

[50] For more on the notion of knowledge-level justification, see §2.2

Gerken's *WA* - that subjects in abnormal cases are *excused*, not justified. It may be that such subjects ought not bear the normative consequences typically associated with transgression of the norm, but they nevertheless engage in an act of wrongdoing. *Pace* Gerken, mere warranted belief, according to proponents of the K-Norm, will *not* do in abnormal cases. In this respect the K-Norm differs fundamentally from the *WA*. The upshot of this is that, firstly, a workable notion of excusability for the K-Norm does not cause it collapse into Gerken's *WA*, and secondly, that proponents of the K-Norm are free to characterise excusability in term of justified belief, or some other kind of knowledge-level-justification-plus-basing condition.

So Gerken's objection raises no problem for the possibility of developing a workable justification/excuse distinction to shore up the K-Norm. A final objection comes from Lackey (2007). Unlike Kvanvig and Gerken, Lackey doesn't attempt to put pressure on the *workability* of the distinction. Rather she argues that it is wrongly applied by proponents of the K-Norm. Lackey asks us to consider our epistemic twins in demon worlds.[51] They share all the same non-factive mental states as those of us in normal worlds who know that p - from the subjective perspective the demon world is indistinguishable from the normal world - but whereas we know that p, their beliefs that p are false, and hence not knowledge. Lackey's argument is simple: she claims that there is a strong intuitive pull to say that such subjects should not be subject to criticism when they employ p in their practical reasoning, and concludes from this that they are fully justified in doing so.

It is hard to know what to make of this objection. Clearly proponents of the K-Norm can, and will, agree with Lackey that it is wrong to criticise subjects in demon worlds for behaving as they do. However, they will deny the assumption Lackey seems to make that non-criticisability is sufficient for justification. Again, they think such behaviour is *excusable*. Lackey is best read, then, as denying that subjects in demon worlds are in any need of excuses, since they are fully justified, and motivating this claim by an appeal to intuition. Essentially she is arguing that proponents of the K-Norm are applying the justification/excuse distinction in the wrong place. What we seem to have then is a clash of intuitions, since, evidently, proponents of the K-Norm do not find it intuitive that demon world subjects are fully justified in employing they false beliefs that p in their practical reasoning. I take it that their reason for this is that in such cases there is one particular dimension on which demon-world subjects beliefs are clearly inferior to their counterparts beliefs: namely, they are *false*.

[51]That is, consider the 'New Evil Demon' scenario described in Cohen (1984).

Given the apparent intuition-clash, it seems fair to say that Lackey's objection is, at best, inconclusive. Ideally we would like something firmer than intuitions to resolve this debate.

§II. Justification Norms

§2.1 Rival Proposals

The criticisms surveyed of the K-Norm's proponent's appeals to the justification/excuse distinction are, variously, unpersuasive and inconclusive. However, the critics have nevertheless taken them, along with the would-be counterexamples, to show that it is theoretically desirable to have an account that obviates the need to appeal to the distinction. With this thought in mind, the critics have proposed rivals epistemic norms of practical reasoning that are, it is claimed, in a better position to capture our judgements about cases.

We have already seen Gerken's (2011) proposal: replace the knowledge condition with a justification condition. A number of other authors offer up variations on this idea. Neta (2009) defends the following principle:

> ***JBK:*** Where S's choice is p-dependent, it is rationally permissible for S to treat the proposition that p as a reason for acting iff S justifiably believes that she knows that p (2009: 686)

Here we are interested in the necessity direction claim of *JBK*. It states that, rather than knowledge, a minimal epistemic constraint on the rational permissibility of a subject treating p as a reason for acting is that the subject *justifiably believes that she knows that p*. This is, in

one sense, a weaker condition than knowledge because a subject may justifiably, *but mistakenly* believe that they know that p. It is compatible with p being false.[52] Neta argues that *JBK* delivers the right verdict about Gettier and justified false belief subjects: since such subjects plausibly justifiably believe that they know that p, they are justified in acting as they do in treating p as a reason for acting. He also argues that it explains all the cases Hawthorne and Stanley use to motivate the K-Norm.

In the context of investigating epistemic norms of assertion, rather than practical reasoning, Lackey (2007) defends a somewhat different J-Norm. It is natural to think that her proposal could be carried over to the case of practical reasoning. The resulting J-Norm would look like this:[53]

> ***RTBNA*:*** It is appropriate for S to employ p in practical reasoning only if (1) It is reasonable for S to believe that p, and (2) If S employed p in practical reasoning, S would do so at least in part because it is reasonable for S to believe that p

The *RTBNA** is to be understood as follows: clause (1) states that S must have a certain degree of propositional justification for the truth of p - enough to make a belief formed by S on the basis of this justification reasonable, where a reasonable belief is one that has justification of sufficient strength to underwrite the subject knowing that p. Clause (2) states that it is necessary that, if S does employ p in practical reasoning, S does so in part *because p is well justified for S*. That is, S's employing p in practical reasoning must be grounded in or based on the justification S has for p. This clause is included to prevent cases where the subject employing p in practical reasoning has excellent evidence for p, but does not pay any heed to this evidence in coming to act as she does, from counting as epistemically appropriate. Clauses like (2) will be important later. Since Gettier and justified false belief cases are ones where, Lackey argues, the subject reasonably believes that p, such cases are perfectly compatible with the *RTBNA**.

---

[52] Though in another sense it may be a stronger condition - if knowledge isn't luminous, then arguably there can be cases where S knows that p but doesn't justifiably believe that she knows that p.

[53] Although it seems likely to me that Lackey would defend the principle I attribute to her here, strictly speaking she is not on the record as arguing for it, since her original arguments only concern epistemic norms of assertion, rather than practical reasoning. The same is true of the principle I will shortly attribute to Kvanvig. In each case I will, for ease of expression, speak as though these authors do indeed defend the epistemic norms of practical reasoning I discuss in this essay. If the reader objects to this, he or she should simply treat the relevant principles as *possible* positions that could be taken in the debate.

Like Lackey (2007), Kvanvig (2010), in the context of assertion rather than practical reasoning, also defends a justification norm. Again it is natural to think that his proposal could be carried over to the practical reasoning case, resulting in the following 'Justification Account':

> ***JA:*** It is appropriate for S to employ p in practical reasoning only if S has epistemic justification for p.

'Epistemic justification' is a term of art for Kvanvig. He takes it, firstly to identify propositional, rather than doxastic, justification, and secondly - like Lackey - to be *justification sufficient to underwrite knowledge.* Understood as such, subjects in Gettier and justified false belief cases have epistemic justification for p. No shortcoming in their justification for p prevents them from knowing that p. Accordingly, *JA* judges such subjects to be justified in acting as they do in employing p as a premise in their practical reasoning.

§2.2. Common Features of The J-Norms

So we have four J-Norms of practical reasoning: *JBK*, *RTBNA**, *WA*, and *JA*.[54] It will be useful for the proceeding argumentation to draw out some important common features amongst them. The first of these is that all the proposed J-Norms should plausibly be understood as requiring, not just some justification, but a certain *level* of justification for appropriate employment of a proposition in practical reasoning: namely *knowledge-level* justification. It can be defined as follows:

> ***K-level Justification:*** S is knowledge-level justified in believing p iff no weakness in

[54]J-Norms of practical reasoning have also been proposed by Littlejohn (2009) and Fantl and McGrath (2009). Littlejohn defends a principle according to which where one's choice is p-dependent, it is appropriate to treat the proposition that p as a reason for acting iff you are justified in believing that p, and p is true. Fantl and McGrath defend the following principle: p is warranted enough to justify you Φ'ing, for any Φ iff you are justified in believing that p. I omit discussion of these proposals partly for reasons of space, and partly because they are not motivated by the claimed undesirability of appeals to the justification/excuse distinction in defense of the K-Norm - the primary concern of this essay. I do think, however, that everything I will go on to say about the proposals that are discussed also applies to these two proposed principles.

p's justification stands in the way of S knowing that p.[55]

Most accounts of knowledge have it that in order for you to know that p, you have, to some degree, to be justified in believing that p. K-level justification is *justification strong enough such that no shortcomings in your strength of justification stand in the way of you knowing that p*. Put intuitively: you are k-level justified in believing that p iff you are justified *enough* to know it. If other factors - truth, belief, absence of Gettier-style luck - conspire in a k-level justified subjects favour, than that subject will know that p, since their justification is strong enough for knowledge.[56] Defined like this, subjects in Gettier cases are k-level justified in believing that p, since it is not their *lack of justification* that prevents them from knowing that p, rather it is the presence of some kind of epistemic luck infecting the relation between their belief and its truth. Were the Gettier-style luck subtracted from their situation, they would in fact *know* that p, and this is true even when there is no change to their level of justification for p. K-level justified belief is also possible for subjects with false beliefs. A subject S has a k-level justified false belief that p just when, were S's belief that p true and not Gettiered, and there were no change to S's level of justification for p, S would know that p.

Kvanvig and Lackey are explicit that the *JA* and *RTBNA** respectively require the subject to have k-level justification. Gerken and Neta don't mention the issue, but the most plausible interpretation of their principles will include such a requirement, since each of the counterexamples to the K-Norm they discuss are one's where the subject plausibly has k-level justification for p, and there are no obvious cases where less-than-k-level justification is sufficient for appropriate employment of a proposition as a premise in practical reasoning; the obvious defectiveness of the lottery reasoning demonstrates this point. It is possible to adapt lottery cases so that's one's evidence is arbitrarily strong, yet not knowledge-level, whilst retaining the judgement that it is inappropriate to rely on the lottery proposition in one's reasoning.[57]

[55]See Fantl and McGrath (2009). Weatherson (personal communication) has expressed to me a concern that k-level justification might be an incoherent notion if it relies on the supposition (which he is sceptical of) that knowledge is 'factorisable'. I won't get into this issue here.

[56] Here is a way of testing whether a subject who doesn't know that p is k-level justified in believing that p: hold fixed their justification, and alter other factors - belief that p, the truth of p, and the absence of Gettier-style luck - affecting knowledge. If you can make it such that they would know that p under certain configurations of external factors, whilst their justification remains fixed, then they are k-level justified in believing p. See Fantl and McGrath (2009) and Conee and Feldman (2004) for more details.

[57] See also Fantl and McGrath (2009)

A second common feature of three of the four proposed J-norms is that they require the subjects employment of p in practical reasoning to be in some way based on, or grounded in, the justification the subject has for p. It is not enough that the subject *has* justification for p, the subject must also employ p in their practical reasoning in part *because* they have this justification. Such a requirement is well motivated, since cases where a subject possess propositional justification for p, but this justification plays no role in their coming to employ p in their practical reasoning, are intuitively cases where the subject is open to (epistemic) criticism for acting as they do. Consider the following case:

> ***TV:*** *Smith is thinking about whether to buy a new plasma screen television. He prefers to do so iff there is at least £1000 in his bank account. In fact there is, and Smith has good evidence for this - he checked his bank account yesterday. However, Smith has temporarily forgotten this fact. Nevertheless, Smith believes he has more than £1000 in his account. This belief is not based on the evidence Smith has, but rather purely on wishful thinking. On the basis of his wishful-thinking-formed belief, Smith proceeds to buy the television.*

If you are like me you will judge that Smith is criticisable for employing 'I have £1000 in my bank account' in his practical reasoning in this case. One should not act on the basis of wishful thinking. This is evidence that Smith has violated a norm. An obvious candidate for that norm is something like: *one's employment of a proposition in one's practical reasoning should be (in some way) sensitive on the evidence one has for the truth of that proposition*. Just what the basing relation will amount to exactly is a difficult question, but it is not one that needs to be addressed here; the notion is, I take it, sufficiently intuitive for our purposes. Roughly speaking, the idea is that it can't be a matter of mere coincidence that the subject employs p in their practical reasoning when they have k-level justification for p; The subject's willingness to employ p in practical reasoning must be somehow sensitive to the facts about their epistemic situation.

I am unable to imagine any circumstances where a subject has good justification for p, but this justification plays no role in the subject coming to employ p in their practical reasoning, where the subject would *not* be criticisable in the way that Smith is. That could just be a failure of imagination on my part, of course. Nevertheless, in the absence of obvious counterexamples I take it that such a condition will be a crucial feature of any plausible J-

Norm of practical reasoning. In this respect, Kvanvig's *JA* proposal is flawed, since it does not include any such basing condition. That doesn't mean *JA* is *false*, however. Indeed, it is entailed by both the K-Norm, and the other J-Norms surveyed. However, in order to refute a claim of the form 'X is a necessary condition for Y', one must show that there is at least one states-of-affairs - Z - that is *sufficient* for Y's obtaining, but where X does not obtain in Z. If it's true that a basing relation is essential for any plausible epistemic norm of practical reasoning, then there will be no cases in which satisfaction of Kvanvig's *JA* will be sufficient for epistemically appropriate practical reasoning (even if necessary), and so he will not have shown that there is no additional, stronger, necessity direction norm of practical reasoning on top of *JA*.[58]

To recap briefly, in this section we have seen that there are two features that any J-Norm must have in order for it to be a plausible rival to the K-Norm. These are (1) the subject must have k-level justification for p, and (2) the subject's employment of p in their practical reasoning must be in some way based on the justification they have for p. In the next two sections I'll explore the question of whether the J-Norms surveyed face similar apparent counterexamples to those facing the K-Norm. I'll argue that J-Norm proponents will likewise have to appeal to the justification/excuse distinction in order to shore up their theories against these counterexamples. I'll then go on to argue that the force of the J-Normers would-be counterexamples to the K-Norm is undermined by this fact.

§2.3 One Potential Way of Counterexampling J-Norms

Each of the three proposed J-Norms can be seen as expressing a variation on the following idea: that one is permitted to employ p as a premise in practical reasoning only if one justifiably believes that p. This idea can be expressed as follows:

> ***JB:*** It is permissible to employ p as a premise in practical reasoning only if you justifiably believe that p

[58] For this reason I will henceforth no longer directly argue against the *JA* norm in the essay.

For the purposes of assessing whether the proposed J-Norms are subject to would-be counterexamples of the kind facing the K-Norm, and so require an appeal to the justification/excuse distinction, it will be useful for us to initially focus on the question of whether *JB* requires an appeal to the distinction, and then, if it does, go on to see whether the considerations that show that it does can be extended to make the same point about the J-Norms proposed by Gerken, Neta, and Lackey.

One apparently promising way of arguing that *JB* is subject to the same kinds of would-be counterexamples to the K-Norm departs from the seemingly platitudinous idea that in general evidence can be misleading and that, as a result it is possible to have false justified beliefs about the answer to just about any question, including the question of whether one is justified in believing that p. As Daniel Greco (2014) puts it: "Justified false beliefs are possible. This claim commands almost universal agreement among contemporary epistemologists. Most would go further and claim that the possibility of justified false belief isn't restricted to special subject matters. Rather, in no domain is justification an infallible guide to truth. At least, if there are domains in which justification entails truth, this requires some special explanation." (2014: 203).

The suggestion, then, would be that since it is in general possible to have misleading evidence, and so justified false beliefs, we should be able to construct cases where one has higher-order evidence (that is, evidence about one's first-order epistemic position) that is sufficiently strong as to justify one in believing that one is justified in believing that p, yet this evidence nevertheless be misleading, and so one is not justified in believing that p. From such cases we could then construct would-be counterexamples to *JB* that are analogous to those that the K-Norm faces, where the subject transgresses the norm but is not properly criticisable for doing so because they had every reason to think that they conformed to it. We would thereby be in possession of cases where proponents of *JB* would have to appeal to the justification/excuse distinction.

Much has been written recently on the impact of higher-order evidence on first-order epistemic states (see for example Christensen 2010, Elga 2013, Horowitz 2013, Lasonen-Aarnio 2014), but most of the literature has focused on cases where the subject does well, epistemically speaking, on the 1st-order, but has higher-order evidence that she has done

badly. So, for example, one focus has been on the question of whether a justified belief that one does not justifiably believe that p can defeat an otherwise justified belief that p. Or interest however is in a different question. What we want to know is whether there can be cases where one does *badly,* epistemically speaking, on the 1st-order, and so one does not justifiably believe that p, but one has higher-order evidence that one has done *well* on the 1st-order, and so one is justified in believing that one does justifiably believe that p. As we may put it, the question is not whether higher-order evidence *defeats* 1st-order justificatory success, but rather whether higher-order evidence *conquers* 1st-order justificatory failure.

If it is in general possible to have misleading evidence, and so justified false beliefs, then the answer to this latter question would seem to be 'no'. And it is not difficult to construct cases that appear to reflect this. Consider the following:

> ***LOGICIAN:*** *Dmitri is a professional logician, and is engaged in some elementary logical reasoning. He arrives at a false conclusion - p - because he affirms the consequent in the course of his reasoning. This mistake comes about because, unbeknownst to himself, Dmitri has recently developed a brain lesion impairing his ability to competently recognise errors in logical reasoning. However, prior to this cognitive impairment, Dmitri was highly competent at logical reasoning, as one would expect from a professional logician. On the basis of his fallacious reasoning, Dmitri comes to believe that p. On the basis of his reasonable belief that he is a highly competent logical reasoner, and the elementary nature of the reasoning he engaged in, Dmitri comes to believe that he justifiably believes that p.*
>
> ***SCIENTIST:*** *Ivan is working on a PhD in biochemistry. He runs a series of experiments to establish whether there is a correlation between the presence of a certain DNA strand and propensity to obesity in mice. Having completed the tests, he comes to believe that there is a correlation and shows his findings to Professor Jones, an expert in the field. Professor Jones studies Ivan's work and, without giving his opinion on the question of whether there is in fact a correlation, tells Ivan that it justifies him in believing that there is a correlation. In fact, Ivan's experiments were riddled with design flaws and fallacious reasoning. Professor Jones didn't pick up on this however, since he was very tired when he studied Ivan's work. Nevertheless, Ivan has no reason to doubt Professor Jones's competence on this occasion, and so comes*

to believe, on the basis of Professor Jones's say so, that he justifiably believes that there is a correlation.

It is very natural to describe the *LOGICIAN* case as follows: Dmitri does not justifiably believe that p. After all, he arrived at p via an egregious logical fallacy (moreover, we can suppose that he knows that affirming the consequent is a logical fallacy). However, despite lacking a justified belief that p, his belief that he justifiably believes that p is itself justified, since the reasoning he engaged in was elementary, he has every reason to believe that he is a highly competent logical reasoner, and he has no reason to suspect that he has made an error in this instance. If this description of the case is correct, then *LOGICIAN* presents a scenario in which the subject does not justifiably believe that p, but justifiably believes that he does justifiably believe that p. The same seems to go for Ivan in *SCIENTIST*. Ivan arrived at his belief that there is a correlation on the basis of an experimental process that was riddled with design flaws and fallacious reasoning. Plausibly such epistemic vices preclude the possibility of Ivan's belief being justified. Nevertheless, since he has been told by an expert whose judgement he has no reason to distrust that his belief is in fact justified, plausibly Ivan justifiably believes that he justifiably believes that there is a correlation. If this description of the case is correct, then again we have a scenario in which the subject does not justifiably believe that p, but justifiably believes that he does justifiably believe that p.

If these descriptions of Dmitri and Ivan are correct, then we can construct would-be counterexamples to *JB* analogous to those facing the K-Norm that require those who would endorse *JB* to appeal to the justification/excuse distinction. Take the case of Dmitri. Suppose that Dmitri goes on to employ p in a piece of practical reasoning as follows:

(1) p is the correct conclusion
(2) If p is the correct conclusion, I shouldn't work on the problem anymore
(3) Therefore: I shouldn't work on the problem anymore.

And suppose that he does so on the grounds of his belief that he justifiably believe that p, and so is permitted, from the epistemic perspective, rely on premise (1). In doing so, he transgresses *JB*, since he does not meet the necessary condition of justifiably believing that which he employs as a premise in practical reasoning. However, quite plausibly it would be unreasonable to criticise Dmitri for doing so, since he had every reason to think that what he

was doing was permissible. Accordingly, we will either have to conclude that *JB* is false - if we don't want to allow for the possibility that there can be blameless violations of it - or accept that, although Dmitri did something impermissible, he is to be excused for his behaviour. Accordingly, just as K-Normers have to appeal to the justification/excuse distinction to handle apparent counterexamples stemming from situations where subjects violate the norm but cannot properly be blamed for doing so, so to will those who endorse *JB,* for precisely the same reason. The same argument can be constructed using Ivan as an example. Suppose that Ivan goes on to employ the proposition that there is a correlation between the presence of the DNA strand and a propensity to obesity in mice as follows:

> (1) There is a correlation between the presence of the DNA strand and a propensity to obesity in mice
> (2) If there is a correlation between the presence of the DNA strand and a propensity to obesity in mice, then I should write an article stating that there is a correlation between the presence of the DNA strand and a propensity to obesity in mice
> (3) Therefore, I should write an article stating that there is a correlation between the presence of the DNA strand and a propensity to obesity in mice

And suppose that he employs premise (1) on the grounds of his belief that he justifiably believes that there is a correlation, and so is permitted, from the epistemic perspective, to do so. In doing so, like Dmitri, he transgresses *JB*, since he does not meet the necessary condition of justifiably believe that which he employs as a premise in practical reasoning. However, as with Dmitri it would be unreasonable to criticise Ivan for doing so, since he had every reason to think that what he was doing was permissible. The same conclusions follow as they do with the case of Dmitri.

Importantly, the cases can be adapted to argue that *JBK, WA,* and *RTBNA** will also have to appeal to the justification/excuse distinction. Consider first *WA*. A reminder, it states that:

> ***WA****:* In the deliberative context, DC, S meets the epistemic conditions on rational use of (her belief that) p as a premise in practical reasoning or of (her belief that) p as a reason for acting only if S is warranted in believing that p to a degree that is adequate relative to DC

Recall, *WA* is best interpreted as requiring that S has, at minimum, k-level justification in believing that p in order to permissibly employ p as a premise in practical reasoning. *LOGICIAN* and *SCIENTIST* both describe cases where the subject does not k-level justifiably believe that p. As a result, they fail to satisfy *WA* when they go on to employ to employ p in their practical reasoning. Yet Dmitri and Ivan are blameless for their violations when they go on to employ the respective propositions as premises in their practical reasoning. Accordingly, we will either have to conclude that *WA* is false - if we don't want to allow for the possibility that there can be blameless violations of it - or accept that, although Dmitri and Ivan did something impermissible, they are to be excused for their behaviour. Accordingly, just as K-Normers have to appeal to the justification/excuse distinction to handle apparent counterexamples stemming from situations where subjects violate the norm but cannot properly be blamed for doing so, so to will those, like Gerken, who endorse *WA,* for precisely the same reason.

Similarly with *RTBNA**. A reminder:

> ***RTBNA*:*** It is appropriate for S to employ p in practical reasoning only if (1) It is reasonable for S to believe that p, and (2) If S employed p in practical reasoning, S would do so at least in part because it is reasonable for S to believe that p

It is not reasonable (in the sense of k-level justification) for either Dmitri or Ivan to believe what they do. By reasoning parallel to that of the last paragraph, we will either have to conclude that *RTBNA** is false, or shore it up with an appeal to the justification/excuse distinction.

Now consider *JBK.* A reminder, it states that:

> ***JBK:*** Where S's choice is p-dependent, it is rationally permissible for S to treat the proposition that p as a reason for acting iff S justifiably believes that she knows that p (2009: 686)

Consider the following adaptation of *LOGICIAN:*

> ***LOGICIAN-JBK:*** *Dmitri is a professional logician, and is engaged in some*

elementary logical reasoning. He arrives at a false conclusion - p - because he affirms the consequent in the course of his reasoning. This mistake comes about because, unbeknownst to himself, Dmitri has recently developed a brain lesion impairing his ability to competently recognise errors in logical reasoning. However, prior to this cognitive impairment, Dmitri was highly competent at logical reasoning, as one would expect from a professional logician. On the basis of his fallacious reasoning, Dmitri comes to believe that p. On the basis of his reasonable belief that he is a highly competent logical reasoner, and the elementary nature of the reasoning he engaged in, Dmitri comes to believe that knows that p.

Dmitri believes himself to know that p. But this belief is not justified. Why not? Because knowledge is factive. If Dmitri did justifiably believe that he knows that p, then he would also justifiably believe that p. Since he doesn't justifiably believe this (just as in *LOGICIAN*), he doesn't justifiably believe that he knows that p. However, suppose that Dmitri also considers whether he justifiably believes that he knows p. Reasonably taking himself to be a highly competent logical reasoner, Dmitri concludes, on the basis of this belief, that his belief that he knows is justified. Plausibly, despite being false, this belief of Dmitri's *is* justified. After all, he has every reason to think that he is a highly competent logical reasoner, and no reason to doubt it, and so every reason to think that he's justified in believing that he knows that p. Now suppose that, as before, Dmitri goes on to employ p in his practical reasoning on the grounds of his belief that he justifiably believe that he knows that p, and so is permitted, from the epistemic perspective, to do so. In doing so, he transgresses *JBK*, since he does not meet the necessary condition of justifiably believing that he knows that which he employs as a premise in practical reasoning. However, it would be unreasonable to criticise Dmitri for doing so, since he had every reason to think that what he was doing was permissible. Accordingly, as with *WA* and *RTBNA** we will either have to conclude that *JBK* is false - if we don't want to allow for the possibility that there can be blameless violations of it - or shore it up with an appeal to the justification/excuse distinction.

§2.4. Concerns about the arguments just given

It is a platitude that evidence can be misleading, and thus that one can have a justified false belief in just about any domain. From this we were able to construct cases, such as *LOGICIAN* and *SCIENTIST*, where it seems right to think that the subjects justifiably but falsely believe that they justifiably believe that p, and so would-be counterexamples to the proposed J-Norms that are analogous to those that the K-Norm faces. These cases appear to show that J-Normers need to appeal to the justification/excuse distinction in order to shore up their theories. However, in this section I will suggest that their persuasive force is questionable, for it may be that J-Normers can acknowledge that, given the platitudes, if there are domains in which justification entails truth, then this is something that requires special explanation, but argue that the domain of justified beliefs about 1st-order justification is a domain with just such a special explanation.

The prospects of arguing that, higher-order evidence aside, Dmitri and Ivan in fact justifiably hold their 1st-order beliefs are bleak. Even internalists will presumably want to maintain that, *ceteris paribus*, beliefs formed on the basis of logical fallacies and flawed experiments aren't justified. However, there may be reasons to think that in cases like *LOGICIAN* and *SCIENTIST*, 2nd-order justification has *downwards generative force*. That is, that the presence of a justified 2nd-order belief that one justifiably believes that p *creates* justification for the otherwise unjustified 1st-order belief. In this section I will briefly sketch a couple of reasons why one might think that. The conclusion I draw is that if we are to be persuaded that there are would-be counterexamples of the kind that require J-Normers to appeal to the justification/excuse distinction, more argumentative machinery will be needed. In §2.5 I provide what I take to be a more compelling argument for the claim.

One reason to think that 2nd-order justification might have downwards generative force is that denying this, and accepting the description of cases like *LOGICIAN* and *SCIENTIST* that was put forward in the last section, commits one to implausible looking claims about what constitutes a rational epistemic agent.[59] Presumably subjects who are justified in believing that p can rationally believe that p, and subjects who are not justified in believing that p can rationally suspend judgement on p or disbelieve that p. If that's right, then insofar as we want to claim that Dmitri and Ivan have 2nd-order justification but not 1st-order justification, we are committed to the claim that Dmitri and Ivan can rationally believe that they are justified in

[59] The point I'm about to make is closely related to a point that Horowitz (2013) makes with respect to cases of higher-order defeat.

believing that p and at the same time rationally suspend judgement on p. But it is far from clear that this combination of attitudes is rationally co-tenable. To see this, imagine having a conversation with Dmitri about his opinions.

> **You**: So, what's the answer then, Dmitri?
> **Dmitri**: I couldn't say
> **You**: Difficult problem is it then?
> **Dmitri**: No, it's easy, and given my calculations, the thing that I ought to believe is that the answer is p
> **You**: Oh, so you think that the answer is p then?
> **Dmitri**: No, I'm not saying that. I ought to believe that the answer is p, but I'm withholding judgement
> **You**: I'm confused, if you think that the right thing for you to believe is that the answer is p, then why don't you believe that?
> **Dmitri:** ...

Dmitri is hard to make sense of in this conversation - he seems to have an irrational combination of attitudes. Surely if he believes that he ought to believe that p, then if he is rational he will believe that p? His attitudes as they stand are akratic, for he doesn't do what he takes himself to have sufficient reason to do. Another way of revealing the oddness of the combination of Dmitri's attitudes is to consider how - according to the description of his case in the last section - it would be rational for him to behave in response to being offered a bet on whether p.[60] If Dmitri is justified in believing that he is justified in believing that p, but not justified in believing that p, and rational subjects doxastic attitudes conform to their justificatory status, then if Dmitri is offered a bet on p which his preferences make it rational for him to take iff he is justified in believing that p, then Dmitri, if rational, will decline the bet whilst at the same time maintaining that he is making a mistake in doing so. Really, he will say, he should be taking the bet. Again, this does not look like the behaviour of a rational person. The general problem is that the claim that one can justifiably but falsely believe that one justifiably believes that p leads to the conclusion that one can rationally believe that one ought to believe that p yet still rationally suspend judgement on whether p or disbelieve p, and these appear to be irrational combinations of attitudes. J-Normers who wish to deny the

[60] Again, this point is an adaptation of a point made by Horowitz (2013)

description of the *LOGICIAN* and *SCIENTIST* cases that lead to the conclusion that they must appeal to the justification/excuse distinction may want to take the apparent irrationality of such combinations of attitudes as a reason to think that despite the general non-factivity of justified belief, in the special case of 2nd-order justified beliefs about first-order justified beliefs, 2nd-order justified beliefs are factive, since they have downwards generative force. If so, then J-Normers may have the resources to deny the argument of the last section purporting to force an appeal to the justification/excuse distinction upon them. At the very least, it seems as though they are in a position to maintain that the *LOGICIAN* and *SCIENTIST* cases don't clearly support that conclusion. More argument will be needed.

A second concern about the argument of the last section stems from the description of the *LOGICIAN-JBK* case that purports to show that Neta's *JBK* theory will require an appeal to the justification/excuse distinction. It was claimed that Dmitri *doesn't* justifiably believe that he knows that p, but *does* justifiably believe that he justifiably believes that he knows that p. The claim that Dmitri doesn't justifiably believe that he knows that p was argued for from the observation that, had he justifiably believed that he knew, then, since knowledge is factive, he would *a fortiori* justifiably believe that p, and his fallacious reasoning precludes the possibility of his belief that p being justified. Given this point about the factivity of knowledge, it is essential for the argument of the last section that Dmitri doesn't justifiably believe that he knows that p, even though he justifiably believes that he justifiably believes that p. Otherwise we wouldn't have would-be counterexamples to *JBK*. But J-Normers might protest that this claim is implausible, since the evidence that justifies Dmitri in believing that he is justified in believing p - namely his evidence that he is an excellent logician, and that the reasoning he engaged in was elementary - also justifies him in believing that he *knows* that p. Such an objection would put pressure on the description of the case offered in the last section, and thus on that section's argument. Moreover, it might also be argued that this state of affairs - one where the subject is justified in believing that they justifiably believe that p, but not justified in believing that they know that p - describes an *impossibility*. According to certain conceptions of justification - conceptions which J-Normers may wish to endorse - justification is best conceived of as the 'appearance of knowledge' (see e.g. Ichikawa 2014, Bird, 2007, and Reynolds, 2013), in the sense that a subject's belief that p is justified only if it is *internally indiscriminable* to the subject from an item of knowledge. If such a theory of justification can be made to work, then it may be that J-Normers can argue that a subject is justified in believing that they justifiably believe that p only if the subject is justified in believing that

they know that p. And in that case, higher-order justification would have downward generative force, as the following argument would show:

> (1) If S justifiably believes that S knows that p, then, *a fortiori*, S justifiably believes that p
>
> (2) If S justifiably believes that S justifiably believes that p, then S justifiably believes that S knows that p
>
> (3) Therefore, if S justifiably believes that S justifiably believes that p, then S justifiably believes that p

So again, it seems that J-Normers may be able to offer a principled reason to resist the description of the *LOGICIAN* and *SCIENTIST* cases put forward in the last section and argue that higher-order justification has downward generative force of a kind that blocks the need for them to appeal to the justification/excuse distinction. So as before, more argument will be needed if we are to accept the claims of the last section purporting to show that J-Normers will need to appeal to the distinction.

§2.5. Anti-Luminosity

More would need to be done to flesh out the details of these lines of argument in order for them to motivate us to reject the argument of §2.3. Nevertheless, it should be clear that there is room for maneuverer here for J-Normers. Thus, if we are to be persuaded that the proposed J-Norms are subject to would-be counterexamples analogous to those facing the K-Norm that force appeals to the justification/excuse distinction, more argument will be needed. In this section I present what I take to be a more compelling argument.

As we saw earlier, Hawthorne and Stanley think that anti-luminosity considerations provide an argument for our desired conclusion. Amia Srinivasan agrees. She remarks that "...If there are no conditions such that we are always in a position to know whether we are in them, then

it is possible to unknowingly, and thus blamelessly, violate any norm" (manuscript)

I think that these authors are right. I am persuaded by the anti-luminosity argument of Williamson (2000) and also that anti-luminosity considerations provided the needed additional argumentative machinery that was lacking. However, Srinivasan and Hawthorne and Stanley don't tell us just how the argument would go, and it is far from clear at first pass how anti-luminosity entails our desired conclusion. To see this, compare the claim of anti-luminosity as applied to transgressions of *JB* to the nature of the would-be counterexamples that the J-Norms would face were they analogous to those that the K-Norm faces. Anti-luminosity claims that:

> ***Anti-luminosity***: For any non-trivial condition C, it is possible for C to obtain but S does not know that C obtains

Applied to the condition of failing to satisfy *JB*, the claim then, is:

> ***JB-Anti-Luminosity***: It is possible for S to fail to satisfy *JB*, but S does not know that S fails to satisfy *JB*

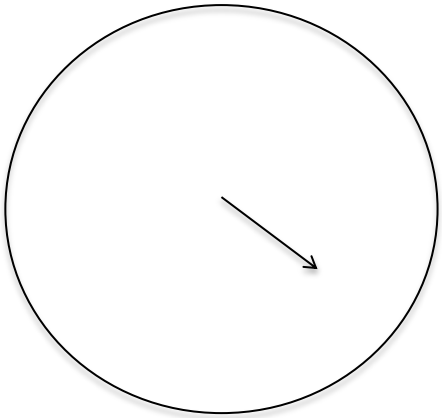
And analogous to the would-be counterexamples to the K-Norm, would-be counterexamples to the proposed J-Norms would have a structure whereby the conditions laid down in the proposed norms are not satisfied, but the subject has every reason to think that they are, and so is justified in believing that they satisfy the norm. So, in the case of *JB*, analogous would-be counterexamples would have a structure whereby the subject fails to satisfy *JB*, but has every reason to think that they do satisfy it, and so is justified in believing that they satisfy it. Does *JB-Anti-Luminosity* entail that there will be cases with this structure? Not obviously. *JB-Anti-Luminosity* only obviously entails that there will be cases where subjects fail to satisfy the conditions laid down in *JB, but are not in a position to know this*. But it is one thing not to *know* that one is transgressing a norm, and quite another to justifiably believe that one is not transgressing it, and, *prima facie* at least, it is the fact that the subjects in would-be counterexamples to the K-Norm justifiably believe that they are acting in a permissible manner that renders them blameless - and thus requires an appeal to the justification/excuse distinction - rather than the fact that they don't know that they are transgressing the norm. Merely *not knowing* that you are transgressing a norm will not necessarily render you

blameless for your transgression, for it may be that, although you don't know that you are transgressing, you have good evidence that you are. And in such a case, it is implausible to think that merely appealing to the fact that you didn't know will get you off the hook. Accordingly, there is no immediately obvious route from luminosity failure to the conclusion those who endorse *JB* must also appeal to the distinction.

Nevertheless, I want to argue, there is a route. It can be found by combining two observations. Firstly, that the denial of luminosity for justification can be pushed beyond the conclusion that a subject can fail to justifiably believe that p without the subject being in a position to know this, to the conclusion that a subject can fail to justifiably believe that p even though it is arbitrarily improbable short of 0 for the subject that this is the case - that is to say, even though it is arbitrarily probable short of 1 for the subject that they do not fail to justifiably believe that p. Secondly, even supposing that anti-luminosity considerations don't show that, for an epistemic state like justified belief, one can be in a situation where that epistemic state does not obtain but one is in a position to justifiably believe that it doesn't obtain - and I will suggest that it is unclear whether anti-luminosity considerations can show this - the observation that *JB* can fail to obtain when it is arbitrarily improbable short of 0 for the subject that it fails to obtains is sufficient to force those who would endorse *JB* into appealing to the justification/excuse distinction.

We can show that it is possible to fail to satisfy *JB* even though it is arbitrarily probable short of 1 for you that you satisfy it by adapting a recent argument from Williamson (2014) to the effect that it is possible to know that p even though it is arbitrarily improbable short of 0 that you know that p. First I'll present Williamson's argument, then I'll show how it can be adapted for our purposes. Williamson asks as to imagine that you are standing some distance from an unmarked clock face with a single hand:[61]

[61] My presentation of the Williamson's argument departs in some respects from his, but nothing important turns on the alterations - they are made merely make the argument clearer.



Suppose that there are 1000 unmarked equally spaced points around the edge of the clock face. Suppose that the hand is in fact pointing at 272. Williamson observes that there is a strongest thing that you can know by visual inspection about where the hand is pointing. You can know some things about the hand's position - for example, that it isn't pointing at 1 or 999, or 100 or 900, but given the limits on your discriminatory capacities, you can't know *exactly* where it is pointing. The strongest thing that you can know is that it is within an interval {-n, 272, +n}, where n = the distance from 272 that, for all you know, the hand could be pointing at. Let us suppose that your discriminatory capacities are such that n = 10. In that case, the strongest thing that you can know about the hand's position is that it is within the interval {262,282}. However, you can only know this if the hand is in fact pointing at 272. If it had been pointing at, say, 269, then, given your discriminatory capacities, the strongest thing you could have known about its position is that it was within the interval {259,279}. Let us suppose that you believe, on the basis of your perceptual experience, that the hand is pointing within the interval {262,282}. *Ex hypothesi* you know that the hand's position is within the interval {262,282}. But, as Williamson points out, *it is very unlikely on your evidence that you know this*. Why? Because you *only* know this if the hand is pointing at 272, but of course for all you know it could be pointing at any of 262,263,...,282 other than 272. And if it was, then you wouldn't know that it is pointing between {262,282}. Since, for all you know it could be pointing at any position between 262 and 282 other than 272, it is highly unlikely on your evidence that it is pointing at 272. If we suppose (as Williamson suggests) that you should give equal credence to the propositions that, for each of 262,...,282 the hand is pointing at 262,...,282, then your credence that it is pointing at 272 should be .05. Accordingly, here we have a case where you know that p (where 'p' = 'The hand is pointing between {262,282}'), but it is highly unlikely on your evidence that you know that p. And we

could decrease the likelihood on your evidence that you know that p arbitrarily low short of 0 by increasing the number of points around the edge of the clock face.

In order to see how this kind of case can be adapted to deliver the conclusion that you can fail to satisfy *JB* even though it is highly likely on your evidence that you satisfy it, the first thing to note is that, as Williamson points out, the conclusion of the clock case can easily be extended from showing that it is possible to have improbable knowledge to showing that it is possible to have improbable justified belief. Just as there is a strongest thing that you know about the hand's position, so there is a strongest thing that you are justified in believing about the hand's position. Suppose that your discriminatory capacities are such that the strongest thing you can justifiably believe about the hand's position is, as before, that it is somewhere between plus +10 and -10 of it's actual position. In that case the strongest thing that you can justifiably believe about the hand's position is that it is within the interval {262,282}. But again, you can only justifiably believe this if the hand is in fact pointing at 272. And for all you know it could be pointing at any of 262,263,...,282. If it was, then you wouldn't justifiably believe that it is pointing between {262,282}. Accordingly, supposing that you believe that it is pointing between {262,282) then here we have a case where you justifiably believe that p (where 'p' = 'The hand is pointing between {262,282}'), but it is highly unlikely on your evidence that you justifiably believe p. And of course, as before, we could decrease the likelihood on your evidence that you justifiably believe that p arbitrarily short low of 0 by increasing the number of points around the edge of the clock face.

The second step towards our desired conclusion that it is possible to fail to satisfy *JB* even though it is highly likely for you that you do satisfy it is to observe that the case can be amended from showing that it is possible to justifiably believe something even though it is highly unlikely for you that you are justified in believing it, to showing that it possible to *fail* to be justified in believing something even though it is highly *likely* for you that you are justified in believing it. To see this, observe that since you are at most justified in believing that the hand is pointing somewhere in the interval {262,282}, you are *not* justified in believing that <the hand is pointing somewhere within an interval with a higher low-end than 262 or a lower high-end than 282> - your discriminatory capacities are not good enough for such a belief to be justified. Call this proposition 'q'. Notice that you are *only* not justified in believing q if the hand is in fact pointing at 272. However, it is of course highly likely on your evidence that the hand is not pointing at 272. It follows that it is highly likely on your

evidence that you are justified in believing q. Let us suppose that you believe q. In that case, we have a situation in which there is a proposition - q - which you unjustifiably believe where it is nevertheless highly likely on your evidence that you do justifiably believe it. More perspicaciously put, the argument is as follows:

(1) If the strongest thing you are justified in believing about the hand's position is that it is within the interval (262,282}, then you are not justified in believing q

(2) The strongest thing you are justified in believing about the hand's position is that it is within the interval (262,282}

(3) Therefore, you are not justified in believing q

(4) You are only not justified in believing q if the hand is pointing at 272

(5) It is highly unlikely on your evidence that the hand is pointing at 272

(6) Therefore, it is highly unlikely on your evidence that you are not justified in believing q

(7) Therefore, it is highly likely on your evidence that you are justified in believing q

(8) Therefore, you are not justified in believing q and it is highly likely on your evidence that you are justified in believing q

So here we have a case in which you do not satisfy the condition - justified belief - laid down by *JB* as necessary for permissibly taking p as a premise in your practical reasoning, but it is highly likely for you that you do. The last step is to imagine that you take q premise in your practical reasoning. Let us suppose that you are offered a bet on whether or not q is true which offers £100 if q is true and loses you £10 if q is false. You reason as follows:

(1) q
(2) If q, and I accept the bet, then I win £100
(3) Therefore, if I accept the bet, I will win £100

(4) I desire to win £100

(5) Therefore, I should accept the bet.

In doing so, you violate *JB* by employing q - which you are not justified in believing - in your practical reasoning. Are you criticisable for doing so? I want to suggest that you are not. Remember, although you are not justified in believing q, not only are you not in a position to know this, but it is highly likely on your evidence that you *are* justified in believing it. We can suppose that it was your awareness of this latter fact that led you to reason as you did. Imagine a third-party holding you to blame for your reasoning. Their complaint is that in relying on q in your reasoning, you transgressed *JB*. In blaming you for your transgression, I suggest that the third-party is being unreasonable. It seems perfectly legitimate for you to respond to the criticism by arguing that you should be excused on the grounds that, not only did you not know that you were transgressing *JB*, but it was highly likely, given your evidence, that you were in fact *not* transgressing it. If so, then we have a case where you blamelessly violated the norm, and thus a case which will force us either to reject *JB* - if we don't want to allow for the possibility of blameless violations of it - or shore it up with an appeal to the justification/excuse distinction.

And just as before, we can adapt this case of applying anti-luminosity considerations to *JB* to applying them to *WA*, *RTBNA**, and *JBK*, for the same argument could be run to show that it is possible to fail to k-level justifiably believe that p even though it is highly likely on your evidence that you do simply by replacing 'the strongest thing you justifiably believe in the clock case with 'the strongest thing you k-level justifiably believe', and to show that it is possible to fail to reasonably believe that you know that p, even though it is highly likely on your evidence that you do reasonably believe that you know by replacing it with 'the strongest thing you justifiably believe that you know'. It should be clear to the reader by now how we can thereby show that those who endorse these theories will have to appeal to the justification/excuse distinction to shore them up.

§2.6. Paradox?

I take the argument of the last section to show that J-Normers will have to appeal to the justification/excuse distinction in the same way that K-Normers do in order to shore-up their views against would-be counterexamples from cases of blameless transgressions. Unlike the argument of §2.4 which relied on an intuitive but, as we saw, controversial, interpretation of the *LOGICIAN* and *SCIENTIST* cases, this argument only relies on a few fairly innocuous assumptions. Namely that the strongest thing you can justifiably believe about the hand's position is that it is within an interval {-n,272,+n}, that you can only justifiably this if it is in fact pointing at 272, and that, given that for all you know the hand could be pointing anywhere within this interval, it is unlikely on your evidence that the hand is pointing at 272. Since I find it hard to see how J-Normers could plausibly deny any of these assumptions, I find it correspondingly hard to see how they can avoid having to appeal to the justification/excuse distinction. I will discuss the implications of this result for the broader debate shortly, but before I do I want to discuss what might appear to be an impending collision course between the arguments of the last two sections. In the last section I argued that reflection on anti-luminosity considerations shows that for the various proposed J-Norms, it is possible to blamelessly transgress them because it is possible to transgress them even though it is highly probable on one's evidence that one satisfies them. In §2.4 I suggested that there are a number of problems with the idea that one can justifiably but falsely believe that one satisfies the conditions laid down in the various J-Norms. You might worry that this gives rise to a paradox. On the one hand, you might think that the clock case shows that it is possible to justifiably but falsely believe that you justifiably believe that p. On the other hand, it looks as though the claim that such a state of affairs is possible forces us to make implausible claims about rationality (because it delivers the verdict that akrasia is rational) and may in fact simply be outright false (because justified belief is internally indiscriminable from knowledge). In that case we would have a paradox, and it is unclear how we might go about resolving it. I am untroubled by this suggestion however, since I think that there are plausible theories of justification that deliver the result that the appearance of paradox here is illusory. The paradox would only arise if we took the clock case to show that it is possible to justifiably but falsely believe that one justifiably believes that p. But closer inspection reveals that this is highly disputable. All that it shows is that it is possible to fail to justifiably believe that p when it is *very likely* on your evidence that you in fact justifiably believe that p. And, as many have taken lottery cases to show, a proposition's being arbitrarily probable on your evidence isn't sufficient to outright justify you in believing it. Notice that in the clock case you, like the lottery ticket holder, but unlike Gettier victim, are in a position to *know* that you

don't know that the hand isn't pointing at 272. Accordingly, you are in a position to know that you don't know that you are justified in believing q. If, as may well be the case, justification is internally indiscriminable from knowledge, then it follows that the clock case does *not* show that it is possible to justifiably but falsely believe that one justifiably believes that p, and in that case the potential paradox can be avoided.

III. Conclusion

A number of commentators have attempted to undermine the claim that knowledge is the norm of practical reasoning by arguing that in certain cases - namely Gettier cases and justified false belief cases - subjects who take propositions that they do not know to be true as premises in their practical reasoning are not to be criticised or blamed for their actions. The critics take this to show that knowledge is not the norm of practical reasoning. The response of K-Normers has been to accept that blame and criticism is inappropriate in these cases, but argue that this is because the subjects are merely to be excused for transgressing the norm, and excusable transgressions of a norm are not counterexamples to that norm. Moreover, they have claimed, any rival norms will be subject to analogous would-be counterexamples. The response of the critics has been to reject the excuse maneuverer and develop rival accounts that deliver the verdict that subjects in Gettier cases and justified false belief cases act permissibly in relying on their beliefs in practical reasoning. The goal of this essay has been to assess the viability of the excuse maneuverer and the K-Normers claim that rival justification norms will be subject to analogous would-be counterexamples. It has been argued that the various objections to the excuse maneuverer are either unpersuasive or inconclusive, and that the rival J-Norms will indeed be subject to analogous would-be counterexamples where subjects blamelessly, and hence excusably, transgress the relevant norms. Where does this leave the debate? An obvious result is that the critics objection to the K-Norm: that the

fact that it can be unwittingly, and hence blamelessly, violated requires that K-Normers make an appeal to the distinction between justifications and excuses that is theoretically problematic, is impotent. There are no good reasons to think that appealing to the distinction in the context of defending theories of the epistemic norms of practical reasoning is problematic, and given that weaker justification norms will also have to appeal to the distinction to handle analogous would-be counterexamples, there is no obvious benefit to be had from developing rival accounts which deliver the verdict that those who act on the misleading appearance of knowledge do so permissibly. However, it would be a mistake to think that this provides a full vindication of the K-Norm. It is one thing to show that no theoretical gain is provided by attempting to develop theories of the epistemic norms of practical reasoning that do not allow for there to be blameless transgressions, and thus avoid the need to appeal to a distinction between justification action and excusable action. It is quite another to show that the correct place to draw the line of the justification/excuse distinction in the case of epistemic norms of practical reasoning is at the level of knowledge, rather than justified belief. Recall, one way of interpreting Lackey's objection to the K-Norm was as claiming that, disagreements about the necessity of justification/excuse distinctions aside, it is simply counterintuitive to suggest that subjects who justifiably believe that p, but don't know that p, do anything that *requires* excusing when they rely on p in their practical reasoning. Although some - myself included - do not find this counterintuitive, it seems likely that just as many will share Lackey's judgement. Accordingly we are faced with a clash of intuitions. In the light of this, if either K-Normers or J-Normers want to persuade us that theirs is the right theory, both sides will need to offer more motivation for their respective views.

5. From Moore's Paradox To The Knowledge Norm Of Belief And Beyond

§I. Knowledge Is The Norm Of Belief

§1.1. Introduction

It is widely accepted that one cannot felicitously assert Moorean conjunctions of the form 'p, & I don't know that p'. It is often argued that this datum motivates the claim that knowledge is the norm of assertion in the sense that the speech act of asserting is governed by a rule stating that one must assert only what one knows.[62]

It is also often claimed that one cannot rationally *outright believe*[63] conjunctions of the form 'p, & I don't know that p'.[64] Can we argue in parallel that knowledge is the norm of belief, in the sense that belief is governed by a rule stating that one must believe only what one knows?

Most epistemologists would reject such an inference.[65] Surely, it will be argued, one is epistemically permitted to believe those propositions it is rational for one to believe? And surely there can be rational beliefs that don't amount to knowledge? The Gettiered, the envatted, and the epistemically unlucky with strong misleading evidence, seem to have rational beliefs. But their beliefs are not items of knowledge. That it is epistemically permissible to believe propositions that one does not know to be true is as widely accepted as

[62] Moore (1962), Hintikka (1962), Unger (1975), Slote (1979) Jones (1991), DeRose (1991, 2002, 2009), Williamson (1996, 2000), Hawthorne (2004), Turri (2011), amongst others

[63] By 'outright belief' I mean a state that involves fully committing oneself to the truth of p, rather than merely taking p to be highly probable. There is, of course, nothing odd about saying 'I believe that p, but I don't know that p'. Plausibly this is because doing so involves hedging to *avoid* committing oneself to p. There is, I take it, nothing irrational about taking p to be highly probable (short of 1) whilst believing that one does not know that p. Thus, I commit myself to rejecting a metaphysics of outright belief where it can be reduced to subjective probability above a threshold <1. For more on these matters, see Huber ed. (2009), as well §2 of this essay. Hereafter when I talk simply of 'belief' I mean outright belief, unless otherwise stated. I intend 'rational' to be read as roughly synonymous with 'reasonable'

[64] Those who argue that it is irrational to believe Moore-paradoxical conjunctions include: Sorensen (1988), Heal (1994), Shoemaker (1994), Williamson (2000), de Almedia (2001), Adler (2002), Sutton (2005, 2007) Huemer (2007), Hill and Schechter (2007), Littlejohn (2010, 2012), Smithies (2012a, 2012b), and Gibbons (2013). (Note that some of these authors only discuss conjunctions of the form 'p, but I don't believe that p'). Not everyone thinks this. For the dissenting view see McGlynn (2014) I won't defend the irrationality of believing Moorean conjunctions here. Rather, I take it as my starting point.

[65] A small but growing number of epistemologists accept the knowledge norm of belief. They include Sutton (2005, 2007), Williamson (2013), and Littlejohn (2013)

the claim that one cannot felicitously assert 'p & I don't know that p'.[66]

In this essay I will argue that, in conjunction with certain plausible assumptions, the impermissibility of believing that p without knowing that p *is entailed by* the irrationality of believing 'p, & I don't know that p'. Given that it is permissible to believe that p if one knows that p, it follows that it is permissible to believe that p iff one knows that p. Knowledge is the norm of belief.[67]

The plan is this: §1.2 presents the master argument. §1.3 motivates its controversial premises and shows how they entail the conclusion. §2 explores the question of *why* it is permissible to believe only what one knows. In §3 a number of objections are considered. §4 concludes.

---

[66] The list of epistemologists who think it is permissible to believe propositions that one does not know to be true is too long to include here. Suffice it to say that it includes the majority of those who work in the field.

[67] It is important to note that in arguing that it is permissible to believe that p iff you know that p, I do *not* intend to suggest that it is rational to believe that p only if you know that p. On the contrary, there are reasons to think that the fact that it is permissible to believe that p iff you know that p *explains why* it is rational for the Gettiered, the envatted and the unlucky to believe that p. As Lasonen-Aarnio (2010) and Williamson (manuscript) have emphasised, despite their transgressing the primary norm of belief, the Gettiered, the envatted, and the unlucky nevertheless do well as believers in an important respect, namely that in believing that p in situations in which it appears as though they know that p, they display a disposition to conform with the knowledge norm. It's just that due to abnormal circumstances the manifestation of this disposition is blocked on the particular occasion of their believing that p. Given that circumstances are normally *not* abnormal, plausibly it is entirely rational for the Gettiered, the envatted, and the unlucky to believe that p despite not knowing that p - in normal worlds, which they may rationally presuppose themselves to be in, their beliefs would be items of knowledge. Moreover, arguably it is rational for them to believe that p precisely *because* it is permissible to believe that p iff one knows that p. Had the conditions under which it is permissible to believe that p not been knowledge, but rather something else, then the conditions under which it is rational to believe that p would have been correspondingly different. Accordingly, the fact that it is permissible to believe that p iff one knows that p may explain why it is rational to believe that p in certain non-knowledge circumstances such as those that the Gettiered, the envatted, and the unlucky, find themselves in. Of course, there is much more to be said here. I don't for a moment suppose that this brief account constitutes a full exposition and defense of a theory of rational belief derived from the knowledge norm. That task is not undertaken here, however, for the main aim of this essay is to argue *for* the knowledge norm of belief and explore the question of *why* knowledge, rather than something else, is the norm. It is worth noting, however, that if the norms of rational belief can be explained by the knowledge norm for permissible belief in the way just suggested, a recent objection to the knowledge norm from Cohen and Comesana (2013a, 2013b, Cohen (manuscript)) is vitiated. Cohen and Comesana argue that those who endorse the claim that a belief is justified iff it is an item of knowledge ('J=K') are using 'justified' in a technical sense. They of course recognise that who endorse J=K are free to use 'justified' in a technical sense and equate it with knowledge, but they demand that if they do, then given that this is a technical sense of justification, they must explain why believers and theorists should care about beliefs being justified in this sense. My own view is that 'justified' admits of several senses, including both a sense that is roughly synonymous with 'rational' or 'reasonable', and a sense that is roughly synonymous with permissible and contrasts with excusability. It should be clear why we (both believers and theorists) should care about the latter sense of justified belief. One of the central goals of epistemology is to try to understand the nature of our epistemic systems - and in particular the conditions under which they permit belief in a proposition. The answer, I will be arguing, is: when, and only when, that belief would be an item of knowledge. Clearly, if to use 'justified belief' to refer to a belief that is epistemically permissible is to use the term in a technical sense, it is nevertheless a sense that epistemologists should take great interest in. Moreover, if Lasonen-Aarnio and Williamson are on the right track, even the project closer to Cohen and Comesana's heart - namely understanding the nature of epistemic rationality - needs to look to this sense of 'justified', for it is then in virtue of the fact that it is permissible to believe that p iff you know that p that rational belief has the character and normative force that it does.

§1.2 The master argument

The master argument is simple:

(1) It is irrational to believe 'p & I don't know that p'

$\sim RB(p \ \& \sim Kp)$

(2) If one knows that p, one rationally believes that p

$Kp \rightarrow RBp$

(3) Therefore, it is irrational to believe that p whilst knowing that one does not know that p

$K\sim Kp \rightarrow \sim RBp$

(4) If it is irrational to believe that p whilst knowing that one does not know that p, then it is not permissible to believe p when one does not know that p

$(K\sim Kp \rightarrow \sim RBp) \rightarrow (\sim Kp \rightarrow \sim PBp)$

(5) Therefore: it is permissible to believe only what one knows.

$PBp \rightarrow Kp$

Premise (1) simply states our starting point: that it is irrational to believe Moorean conjunctions. Premise (2) is almost universally accepted by epistemologists. (3) follows from (1) and (2) as follows:

(a) It is irrational to believe 'p, & I don't know that p'

~RB(p & ~Kp)
- premise (1)

(b) If one knows that p, one rationally believes that p

Kp → RBp
- premise (2)

(c) One knows that one does not know that p, and one rationally believes that p

K~Kp & RBp
- assume for reductio

(d) Therefore: one rationally believes that one does not know that p

RB~Kp
- b,c

(e) One rationally believes that p

RBp
- c, &E

(f) Therefore: one rationally believes 'p, & I don't know that p'

RB(p & ~Kp)
- d,e, &closure

(g) (f) contradicts (a).

So we must either reject (1) or (2) or accept (3): that it is irrational to believe that p whilst knowing that you don't know that p. Since we should accept (1) and (2), we should accept

(3).[68]

(5) is our conclusion, that knowledge is the norm of belief. But why accept premise (4)? It is far from obvious that the irrationality of believing p whilst knowing that one does not know that p entails that that it is impermissible to believe that p when one does not know that p. This needs to be argued for. It is to this task that we now turn.

§1.3. Permission, culpability, and excusability

§1.3.1 The gap that excuses fill

The aim of this section is to argue that the impermissibility of believing that p without knowing that p is entailed by the irrationality of believing that p whilst knowing that one does not know that p. That's what premise (4) states. I motivate this claim by arguing that it follows from the nature of the relationship between permissibility, culpability, and excusability.

Roughly put, the argument will be this: it follows from the fact that it is irrational to believe that p whilst knowing that you don't know that p, that it can only be rational to believe that p whilst not knowing that p *when you are unaware that this is what you are doing* (that is, when you don't know that you don't know that p). This fact, I argue, implies that believing that p without knowing that p is impermissible, because it is only when you have an *excuse* for believing that p without knowing that p - the excuse being that you did not know that you were doing this - that it can be rational for you to do so. And if it were permissible to believe that p without knowing that p, you wouldn't need an excuse in order to rationally do so. This section is dedicated to explicating and motivating this claim.

We begin by considering the relationship between permissibility, culpability, and excusability. A very natural account of the relationship goes as follows. If Φ-ing is impermissible, then if S Φ's, S does something wrong. Often when you do something wrong, you are culpable for your

[68] The derivation of (f) from (d) and (e) assumes that rational belief is closed under conjunction. You might think that rational belief isn't closed under conjunction, in which case you might have qualms about this argument. I address these concerns in §3.

behaviour. But not always; sometimes you have an *excuse*. Excuses diminish or remove your culpability for wrongful behaviour without also turning it into rightful behaviour. Excuses have two important properties. Firstly, a fact F can only excuse you for Φ-ing if you were not permitted to Φ. If you were permitted to Φ, then not only do you not *need* an excuse, you cannot *have* an excuse, since having an excuse for Φ-ing entails that Φ-ing was impermissible. Hence the strangeness of 'S did nothing wrong, but we excuse him for it'. If S did nothing wrong, one is lead to wonder, what is he being excused for? Only wrongdoing calls for an excuse. Secondly, excuses exhaustively fill the gap between wrongdoing and culpability. If you Φ, and Φ-ing is impermissible, then if you don't have an excuse for Φ-ing, you are culpable for Φ-ing. Hence the strangeness of 'S shouldn't have Φ'ed, but he did anyway, and he doesn't have any sort of excuse. Nevertheless, he's not to blame'. If S did something wrong, and he doesn't have an excuse, surely he *must* be blameworthy? If you wrong, you need an excuse to get off the hook.[69]

If this view of excuses is right, then it entails the following principle:

> ***1st Permissibility principle****:* Φ-ing is permissible for S iff it is possible for S to non-culpably Φ without an excuse.[70]

Not only does this principle follow from a very natural account of the relationship between permissibility, culpability, and excusability, it also makes good intuitive sense. If you are permitted to Φ, then you shouldn't need an excuse to Φ without being held culpable. You only need an excuse - and moreover, as we just saw, can only *have* an excuse - when you do something wrong, and you, *ex hypothesi,* didn't do anything wrong. On the other hand, if the only way you can avoid being culpable for your behaviour is by having an excuse, then it must be that you were not permitted to do what you did.

It will be useful to look at a case to get a feel for this principle. So consider the following

[69] That's not to say that excuses are the only thing that diminishes culpability. Sometimes people are so far gone that it is not appropriate to hold them to the standards others are held to. In that case, they have an exemption. But if they are not appropriately held to the relevant standard, then they cease to have an obligation to conform to it. So exemptions don't threaten the claim that excuses exhaustively fill the gap between wrongdoing and culpability. (cf. Gardner (1997))

[70] Assuming that Φ-ing is impermissible iff not Φ-ing is obligatory, an equivalent principle is:

> *1st Obligation principle:* not Φ-ing is obligatory for S iff it is not possible for S to non-culpably Φ without an excuse.

scenario, borrowed from Elizabeth Harman (2011):

> Cyanide Coffee
>
> *Anne spoons some cyanide into the coffee she is making for Bill. Bill drinks the coffee and dies. Given suitable background assumptions (for example, Anne was not acting in self defense), Anne did something morally impermissible. However, she is not culpable for her behaviour, since she reasonably believed that the cyanide was in fact sugar. It turns out that Bill's nemesis surreptitiously switched the contents of the sugar bowl when no one was around.*

Anne did something impermissible in poisoning Bill, but she is not culpable because she has an excellent excuse - she was blamelessly unaware of what she was doing. Given that what she did was impermissible, the principle states that it is not possible for Anne to be non-culpable without an excuse. That seems right. Had Anne known what she was doing, and not had any other kind of excuse, she would surely have been culpable for poisoning Bill. On the other hand, had there not been any malevolent interference, and so Anne had simply spooned some sugar into Bill's coffee, rather than cyanide, then she would have had no need for an excuse for her behaviour, since she would have done nothing that required excusing, and it is very odd to think that she would need something that would normally serve to excuse to be in place in order not to be culpable for what she did.[71]

The *1st permissibility principle* provides us with a test. We can establish whether Φ-ing is permissible or not for S by establishing whether it is possible for S to non-culpably Φ without an excuse. If it is possible, then S is permitted to Φ. If it isn't, then S is not permitted to Φ.

In order to run such a test, we need, of course, to know under what conditions it is possible to be excused for wrongdoing - that is, what kinds of considerations excuse. Providing an exhaustive list is far beyond the scope of this essay, and unnecessary in any case. But it is often thought that at least the following can serve to excuse (see, e.g. Gardner (1997)): *Duress, Provocation, Intoxication*, and *Misapprehension.* Intuitive cases are easy to come by:

[71] I don't mean to suggest that every case will be this clear-cut. There are tricky cases where Φ-ing is permissible, but ill advised. In such cases, if you Φ it may be that you are properly criticisable. There are also cases where Φ-ing is permitted but all your evidence suggests that it isn't. Again, here you may be properly criticisable, because reckless, if you Φ. These matters will be discussed further in §3.

***Duress**:* If Dan is coerced into smuggling weapons by a gang who have made threats to his wellbeing, then whilst, we may suppose, the threats didn't made his action permissible, the duress may well diminish his culpability. He did something impermissible, but he has an excuse - he was under duress.

***Provocation**:* If Joan lashes out physically at her friend after months of taunting, then whilst, we may suppose, she was wrong to do so, the fact that she was provoked may well diminish her culpability. She wasn't permitted to lash out, but she has an excuse - she was provoked.

***Intoxication**:* If someone slips a pill into Sarah's drink, which causes her to drive dangerously, then although she did something wrong in driving in that manner, it may well be that her culpability is diminished because she was drugged. She wasn't permitted to drive in that manner, but she has an excuse - she was drugged.

Misapprehension: We have already seen a case of misapprehension excusing - Anne-the-poisoner. Here's a similar case: If Tom, a policeman, unwittingly strays on to a movie set and shoots and kills what he reasonably but mistakenly believes to be a fleeing armed robber, who is in fact an actor, then whilst he wronged - he killed an innocent man, after all - the fact that it was reasonable for him to believe that his behaviour was justified may well diminish his culpability. He wasn't permitted to shoot the innocent man, but he has an excuse - it appeared as though he was in a situation where shooting was permitted.

To be clear, the claim here is not that duress, provocation, intoxication and misapprehension *always* excuse, or that they always *fully* remove culpability. If the duress or provocation was very mild, and the wrongdoing very great, then the fact that there was duress or provocation may only partially remove culpability, or may not remove it at all. Likewise, if Sarah is responsible for her own intoxication - she voluntarily drank far too much liquor - the fact that she was intoxicated may do little to nothing to excuse her wrongdoing. Similar considerations apply to other excusing conditions.

§1.3.2. When misapprehension excuses

Given the goal of this section - to argue that if it is irrational to believe that p whilst knowing that one does not know that p, then it is not permissible to believe p when one does not know that p - it should be clear that our interest is in misapprehension cases. That is, cases where appearances diverge from reality in such a way that a subject's wrongful behaviour is excusable. There are a number of questions we can ask about the conditions under which misapprehension excuses. For example, we will want to know what character misapprehension must have in order to excuse. Clearly, as with duress, provocation, and intoxication, it doesn't always. If, rather than straying on to a movie set, Tom was instead walking down the street and shot a young black man, racistly believing that the man was an armed robber, without a shred of evidence, then the fact that Tom believed that the man was a armed robber will likely do very little to diminish his culpability. Mere mistaken belief isn't sufficient for misapprehension to excuse. What else is required then? The question is a difficult one. However, for the purposes of motivating premise (4), we do not need to attempt to answer this question. It will suffice to note that at least *sometimes*, not knowing that you are Φ-ing excuses wrongful Φ-ing, even if it doesn't always. The cases of Anne-the-poisoner and Tom-the-killer-cop show just this.

I want to argue that, in conjunction with the fact that not knowing what you are doing can serve to excuse, the *1st permissibility* principle motivates premise (4).

Now, it might seem as though there is a particularly straightforward way of arguing for this. The argument doesn't work, but as we will see, its failure is instructive. It departs from the thought that, if not knowing what you are doing excuses (at least sometimes), and Φ-ing is permissible for S iff it is possible for S to non-culpably not-Φ without an excuse, then we should be able to cash out the *1st permissibility principle* as follows:

> ***2nd Permissibility principle****:* Φ-ing is permissible for S iff were S to know that she was Φ-ing, then it would be possible for S to non-culpably Φ without an excuse

The thought is this: the *1st permissibility principle* states that Φ-ing is permissible for S iff it is metaphysically possible for S to non-culpably Φ without an excuse. As we have seen, not knowing you are Φ-ing (sometimes) excuses. So it would seem to follow that Φ-ing is permissible iff it is possible for S to non-culpably Φ whilst not having the excuse of not

knowing that she is Φ-ing, and not having any other excuse. And that entails *knowing* that she is Φ-ing and not having any other excuse.

From the *2nd permissibility principle*, it is straightforward to derive premise (4).

> (1*) By the *2nd permissibility principle* if believing that p without knowing that p is permissible, then it is possible to non-culpably believe that p whilst knowing that you don't know that p, without an excuse
>
> (2*) It is irrational to believe that p whilst knowing that you don't know that p
>
> (3*) If you have no excuse for your irrationality, then you are culpable for it
>
> (4*) Therefore: it is not possible to non-culpably believe that p whilst knowing that you don't know that p, without an excuse
>
> (5*) Therefore: it is not permissible to believe that p when you don't know that p

(1*) applies the *2nd permissibility principle* to the case of believing that p without knowing that p. (2*) is the antecedent of premise (4). (3*) will, I take it, be uncontroversial if the *1st permissibility principle* is correct, since it is surely not permitted, from the epistemic perspective, to believe in an epistemically irrational manner. (4*) follows from (1*), (2*), and (3*), and (5*) - which is the consequent of premise (4) - follows from (1*) and (4*).

Thus, if the *2nd permissibility principle* is correct, then, given the uncontroversial assumption that it is impermissible, from the epistemic perspective, to believe irrationally, premise (4) of the master argument holds, and we can simply run the master argument to the conclusion that it is permissible to believe only what one knows.

§1.3.3. When ignorance justifies

Alas, arguing for premise (4) isn't that straightforward. The *2nd permissibility principle* is false. Sometimes not knowing you are Φ-ing doesn't excuse - as it does for Anne-the-poisoner and Tom-the-killer-cop - but rather *justifies*, at least in the restricted sense that there are many

cases in which you must act on the balance of probabilities even though you don't know what the objectively right thing to do is, but were you to have known more, you would have been culpable for acting in the way that you did. To see this, consider a version of Parfit's (2011) miners case:

Miners case

There are two groups of miners trapped down two different mines - mine A and mine B. You only have the resources to save one group. If you dedicate your resources to A, then the miners in A will be saved and the miners in B will perish. Vice versa if you dedicate your resources to B. You know that one mine has 5 miners in it, and the other 100. Unfortunately you don't know which is which. However, your total evidence makes it rational for you to have a .95 credence that the 100 are in mine A, and this is the credence you have. Time is running out, all the available evidence is in, and if you don't act soon, both sets of miners will perish. What should you do?

Plausibly, you are permitted - indeed, obligated - to dedicate your resources to mine A. And this remains true even if the 100 are in fact in mine B. Suppose that the 100 are indeed in B. In such a case, you are permitted to Φ, where Φ-ing = saving the 5 rather than the 100. Notice however, that had you *known* that what you were doing was saving the 5 rather than the 100, you would have been culpable for you behaviour.

Cases like this can be multiplied. What they bring out is a point long recognised by decision theorists and moral philosophers: that what you ought to do is often partly a function of your epistemic position. If, instead, your total evidence had made it rational for you to have a .95 credence that the 100 were in mine *B*, then you would not have been permitted to dedicate your resources to mine A. Rather, you would be obligated to dedicate them to mine B. In an uncertain world, often the right thing to do is to go with what the balance of evidence suggests is the objectively right thing, even if your evidence doesn't put you in a position to know that it is the objectively right thing to do, and even if that evidence is misleading.

Given this fact, the *2nd permissibility principle* is too crude, since it delivers the verdict in the miners case that you are not permitted to dedicate your resources to mine A. That's the wrong result. What is the defect in the principle that leads it astray in cases like this one? The answer

is, I think, quite clear. The problem is that the principle does not allow for changes in your epistemic position to affect what it is permissible for you to do. It doesn't take into account the fact that often an action that is permissible given your epistemic circumstances would not be permissible in different epistemic circumstances. Since changes in epistemic circumstances manifestly can affect the permissibility of a given action, the principle must be altered to take this into account if we are to make sense of the conditions under which ignorance[72] excuses the impermissible, rather than creating a permission because one must act on the balance of probabilities. What we seek is a principle that honors the fact that Φ-ing is permissible iff one can non-culpably Φ without an excuse, and honors the fact that ignorance sometimes functions to excuse, whilst at the same time honoring the fact that changes in epistemic position can affect what actions are permissible. That is, we are looking for a principled different between cases in which ignorance justifies, like the miners case, and cases where it merely excuses, like the case of Anne-the-poisoner and Tom-the-killer-cop.

What's the difference between cases where ignorance justifies, because you must act on the balance of probabilities, and cases where it merely excuses? The important difference, it seems to me, is that in cases where ignorance justifies, like the miners case, you can permissibly save the 5, rather than the 100, *whilst knowing that it is possible that you are doing just this*, without your needing an excuse in order to not be culpable for your behaviour. That is to say: whilst knowing that you don't *know* that you are doing the objectively right thing.[73] Not so for Anne-the-poisoner and Tom-the-killer-cop. Had Anne known that she didn't know that there wasn't cyanide in the bowl, then she would be culpable for proceeding to spoon the substance into Bill's coffee without checking first, unless she had some kind of excuse. Similarly, had Tom known that he didn't know that he wasn't shooting an innocent man - that is, had he known that he didn't know whether or not the man was a genuine armed robber, rather than an actor - he would (other excusing conditions aside) have been culpable for his behaviour. This difference suggests that the right way to adapt the principle in order to reflect the distinction between cases where ignorance justifies because one must act on the balance of probabilities, and cases where ignorance merely excuses because merely having the balance of probabilities isn't enough, is as follows:

[72] Here, and throughout, by 'ignorance' I mean not knowing. So S is ignorant of whether p iff S does not know whether p.
[73] In this section we understand epistemic possibility as being determined by a subject's knowledge. So it is possible for S that ~p iff S does not know that p.

3rd Permissibility principle: Φ-ing without knowing that Φ-ing is objectively right is permissible for S iff were S to know that she did not know that Φ-ing is objectively right, it would nevertheless be possible for S to non-culpably Φ without an excuse

This principle is somewhat difficult to parse, so let me state clearly what it says:

It is permissible for S to Φ when S does not know that Φ-ing is objectively right iff the following subjunctive conditional is true:

Were S to know that S does not know that Φ-ing is objectively right,
then it would be possible for (1)-(3) to be jointly true:

(1) S Φ's
(2) S is not culpable for Φ-ing
(3) S has no excuse for Φ-ing

To illustrate, we can look at what the principle says about the miners case. Recall, the rational credence for you to have given your total evidence is that it is .95 likely that the 100 miners are in mine A, and the 5 miners in mine B, and this is the credence you have. So you don't know that dedicating your resources to mine A is the objectively right thing to do; for all you know the objectively right thing to do is to dedicate your resources to mine B. The question we are interested in is this: in these uncertain circumstances, where you don't know what the objectively right thing to do is, can you be permitted to dedicate your resources to mine A? The *3rd permissibility principle* says that the answer is 'yes' iff it is possible for (a)-(c) to be jointly true when you know that you don't know that dedicating your resources to mine A is the objectively right thing to do

(a) You dedicate your resources to mine A

(b) There is nothing that would serve to excuse you for dedicating your resources to mine A were it impermissible to do so.

(c) You are not culpable for dedicating your resources to mine A.

If it is possible for (a)-(c) to be jointly true in these circumstances, then you are permitted to dedicate your resources to mine A without knowing that this is the objectively right thing to do. If it is impossible for all of (a)-(c) to be jointly true, then you are not permitted.

Why think that the *3rd permissibility principle* is true? As I see it, there are (at least) two reasons. Firstly, the principle delivers the correct verdict in the miners case, and, it seems to me, all relevantly similar cases - that is, cases where you permitted to Φ despite the fact that you do not know that Φ-ing is the objectively right thing to do. As we saw above, In the miners case we get the correct verdict - that you are permitted to save the 5, rather than the 100 - since it is possible for you to non-culpably save the 5, rather than the 100 whilst *knowing* that it is possible that you are doing just this, without any other excusing conditions - intoxication, duress, etc. - being in place. The situation you face in the miners case is a common one: you must act on the balance of the probabilities, in the face of the possibility that you are not doing the objectively right thing. In such situations, it seems to me, it is always possible for you to act without an excuse, without lapsing into culpability. And this includes the excuse of not knowing the risks. Thus, I see no obvious counterexamples to the principles from the direction of cases where ignorance justifies. That is, I see no obvious cases where intuitively Φ-ing without knowing that Φ-ing is the objectively right thing to do is permissible, but it is impossible for conditions (a)-(c) above to be jointly satisfied.

Moreover, it seems to me that there are no obvious counterexamples to be found from cases where ignorance excuses. As we saw above, the obvious difference between you in the miners case, and the cases of Anne-the-poisoner and Tom-the-killer-cop is that, whilst you can permissibly, non-culpably, dedicate your resources to mine A whilst *knowing* that it is possible, given your epistemic circumstances, that you will thereby be saving the 5, rather than the 100, had Tom known that he didn't know whether or not he was shooting an innocent man, and had Anne known that she didn't know whether she was spooning cyanide or sugar into Bill's coffee, Tom and Anne would (other excuses aside) have been culpable for the behaviour. In these cases, having the balance of probabilities isn't good enough. The only reason that Anne and Tom aren't culpable is because they are blamelessly ignorant of their epistemic circumstances. In that case, their lack of knowledge only serves to excuse, rather than justify, their actions. Plausibly, the same is true of other such cases where ignorance excuses, rather than justifies. So, I see no obvious counterexamples to the principles from this

direction either. That is, I see no obvious cases where intuitively Φ-ing without knowing that Φ-ing is objectively right is not permissible for S, but it is possible for conditions (a)-(c) to be jointly satisfied.[74]

The absence of obvious counterexamples to the principles either from cases where ignorance justifies, or from cases where ignorance excuses, is a reason to accept the *3rd permissibility principle*. A defeasible reason, but a reason nonetheless.

The second motivation for the principle is that there are theoretical reasons to expect it to be true. To see this, recall that, as we have already seen, if excuses exhaustively fill the gap between wrongdoing and culpability, then if Φ-ing is permissible, it must be possible to non-culpably Φ without having an excuse. That's what the *1st permissibility principle* states. Now, sometimes you don't know what the objectively right thing to do is, but you must act, and so you are permitted to act on the balance of probabilities. This is what's going on in the miners case, and similar cases. In such cases, what you are permitted to do is partly a function of your epistemic circumstances - what things you know, what things you are ignorant of, what credences your total evidence makes rational, and so on. But once the issue of what you are permitted to do given your epistemic circumstances has been settled - that is, once your epistemic circumstances have been 'factored in' to the set of considerations that determine whether it is permissible for you to Φ - then, given that excuses exhaustively fill the gap between wrongdoing and culpability, it should be possible for you to proceed to do the permissible thing non-culpably without the need for an excuse. And that means without, amongst other things, your non-culpability being a result of your ignorance of your epistemic circumstances.

Why? Because if you are permitted/not permitted to Φ without knowing that Φ-ing is the objectively right thing to do, then considerations about what epistemic circumstances have to obtain in order for you to be permitted/not permitted to Φ *have already been taken into account*, and it has been determined that you are permitted/not permitted to Φ without knowing that Φ-ing is the objectively right thing to do. Given this, if it turned out that the only reason you were non-culpable for Φ-ing without knowing that Φ-ing is objectively right is because of a particular kind of ignorance on your part - namely ignorance of your epistemic

[74] You might think that, provided that it was sufficiently probable on Anne and Tom's evidence that they were doing the right thing, they could non-culpably act as they do without an excuse. I will address this concern in §3.

circumstances (as is the case with Anne and Tom) - then it must be that your epistemic circumstances weren't circumstances in which Φ-ing without knowing that Φ-ing is the objectively right thing to do was permitted in the first place, and the best you can hope to get from your ignorance is an excuse. But by the *1st obligation principle* and the *1st permissibly principle*, if you are permitted to Φ, then it is possible to non-culpably Φ *without* an excuse. Therefore: it must be possible to non-culpably Φ without knowing that Φ-ing is the objectively right thing to do whilst knowing that you don't know whether Φ-ing is the objectively right thing to do. And this is just what the *3rd permissibility principle* states.

From the epistemic perspective, believing that p is the objectively right thing to do if p is true, and the objectively wrong thing to do if p is false (as it is often noted: a belief is correct iff true). Given this, with the *3rd permissibility principle* in hand we are now in a position to motivate premise (4) of the master argument. The argument is parallel to the one given in §1.3.2, but with the *3rd permissibility principle* replacing the flawed *2nd permissibility principle*:

> (1**) By the *3rd permissibility principle,* if believing that p is true without knowing that p is true is permissible, then it is possible to non-culpably believe that p is true whilst knowing that you don't know that p is true, without an excuse
>
> (2**) It is irrational to believe that p whilst knowing that you don't know that p
>
> (3**) If you have no excuse for your irrationality, then you are culpable for it
>
> (4**) Therefore: it is not possible to non-culpably believe that p whilst knowing that you don't know that p, without an excuse
>
> (5**) Therefore: it is not permissible to believe that p when you don't know that p

(1**) applies the *3rd permissibility principle* to the case of believing that p whilst not knowing that p, where believing that p is objectively right iff p is true. (2**) is the antecedent of premise (4). (3**) is uncontroversial if the *1st permissibility principle* is correct, since it is surely not permitted, from the epistemic perspective, to believe in an epistemically irrational manner. (4**) follows from (1**), (2**), and (3**), and (5**) - which is the consequent of

premise (4) - follows from (1**) and (4**).

With premise (4) in hand, we are now in a position to run the master argument:

(1) It is irrational to believe 'p & I don't know that p'[75]

(2) If one knows that p, one rationally believes that p

(3) Therefore, it is irrational to believe that p whilst knowing that one does not know that p

(4) If it is irrational to believe that p whilst knowing that one does not know that p, then it is not permissible to believe p when one does not know that p

(5) Therefore: it is permissible to believe only what one knows.

So if it is irrational to believe 'p, but I don't know that p', then knowledge is the norm of belief in the sense that it is permissible to believe that p only if one knows that p. Since the claim that knowledge that p is sufficient for permissible belief that p is almost universally accepted by epistemologists, this can uncontroversially be strengthened to a biconditional:

Knowledge norm of belief: One is permitted to believe that p iff one knows that p.

[75] You might think that, just as it is irrational to believe 'p, & I don't know that p', it is also irrational to believe 'p, & I can't be certain that p'. If you think that, you might worry that this kind of argument overgenerates, since then an argument parallel to this one could be run to show that one is permitted to believe that p only if p is certain - and this is too strong. I address this concern, and related concerns, in §3.

§II. Why Is Knowledge The Norm Of Belief?

§2.1. The 'why?' question

In §1 I argued that if it is irrational to believe conjunctions of the form 'p, and I don't know that p', the knowledge is the norm of belief in the sense that it is epistemically permissible to believe that p iff you know that p. Since I think that it *is* irrational to believe conjunctions of the form 'p, and I don't know that p', I think that knowledge is indeed the norm of belief.

Here I address a further question. *Why* is knowledge is the norm of belief? Why knowledge, rather than something else? The question is an interesting one for at least two reasons. Firstly, and most obviously, because an answer to it will deepen our understanding of the epistemic; for epistemologists this is a question worth exploring for its own sake. Secondly, because supposing that, as was suggested at the outset of this essay, the fact that knowledge is the norm of belief explains why it is rational to believe that p, when it is, then answering the question will also give us a more complete understanding of rational belief.

One might think that it is simply a brute fact about our epistemic system that knowledge is the norm of belief. I think that's rather implausible. It is natural to think that we have *this* norm, rather than some other norm, of permissible belief because conforming to a normative system in which it is permissible to believe that p iff one knows that p serves our needs and interests in some way. Now, this might ultimately turn out not to be the case; perhaps the knowledge norm serves no particular need or interest of ours. But the hypothesis is one that is worth investigating.[76]

I think we can say some plausible things about why knowledge - rather than something else -

[76] The project of this section shares some affinities with Edward Craig's project in 'Knowledge and the State of Nature' (1990) insofar as it seeks to shed light on epistemological issues by considering the general needs and interests of human beings. The projects differ in some respect however. For example, whereas Craig is primarily interested in why we have the concept KNOWLEDGE, I am more interested in why we care about believing only what we know.

is the norm of belief. In this section I sketch out an answer to the 'why?' question. However, no pretense is made here that the account offered is either exhaustive or definitive. The account put forward here is perfectly compatible with there being additional reasons why knowledge is the norm of belief beyond those suggested, and much of what is said in this section is put forward in the spirit of speculative conjecture, rather than conclusive argument. The reader is asked to bear these facts in mind when reading what follows. This is only a tentative proposal.

Of course, just how one will wish to answer the 'why?' question depends to a significant extent on one's views on the nature of knowledge. My preferences are for a *fallibilist* conception of knowledge, whereby S can know that p even though S could have been wrong about p, and a *safety* conception of knowledge whereby knowledge is simply safe belief - that is, S knows that p iff S could not have *easily* been wrong about p. I will have more to say about these conditions shortly. Both, of course, are controversial (particularly the safety conception). I won't argue for them here, for my goal in this section is to offer an account of how one might explain why knowledge is the norm of belief *given* these assumptions about the nature of knowledge. I do so without apology; conditional inquiry is a valuable enterprise even though its conclusions are not categorical.

§2.2. An Overview

I begin with an overview of the account and then go on to spell out and motivate some of the details.

In slogan form, my suggestion is that knowledge is the norm of belief because a belief is fit to be relied on in decision making when, and only when, it is an item of knowledge. More specifically the thought (roughly stated) is this. If knowledge is fallible safe belief, then S knows that p iff S could not have easily have falsely believed that p, but it is not required that there was no possibility of S falsely believing that p. Belief is a functional kind that has constitutive connections with decision making dispositions. In particular, it is a necessary condition on S believing p that S is disposed to treat p as certain for the purposes of decision

making in normal circumstances. Due to cognitive and epistemic limitations, we have a need to employ a heuristic of treating propositions as epistemically certain for decision making when they are in fact epistemically uncertain. Taking uncertainties as certain for decision making introduces a kind of epistemic risk to the actions that result from such decision making. Lest the epistemic norms of belief license foreseeably unhelpful decision making procedures, we should expect there to be restrictions on when it permissible to take an uncertainty as certain for the purposes of decision making. Conforming to a rule that states that one is permitted to believe that p iff one knows that p is, I want to suggest, a good policy to have given these needs and limitations. And this explains (at least in part) why we have a rule that it is permissible to believe that p iff one knows that p.

Obviously there is a lot to elucidate and motivate here. We will begin with some elucidation of the assumptions and then move on to so see how the account might be motivated given them. There is no scope here to argue for the central assumptions here. For example, I will take it as a given that knowledge is fallible safe belief, that belief is a functional kind with the kind of constitutive connections to action described, and that we employ the heuristic of taking uncertainties as certain in decision making. Arguing for these claims here would only amount to rehearsing already familiar points from debates in epistemology, decision theory, and the philosophy of mind. Instead, the reader will be directed to the work of others in which these claims have been argued for. What I *will* argue is that when they are put together, these assumptions offer us a promising explanation of why knowledge is the norm of belief.

§2.3. Some Assumptions

We begin with fallibilism. There are a number of way to conceive of the fallibility of knowledge. Some (e.g. Fantl & McGrath 2009, et al) do so in terms of epistemic probabilities. The claim being that S can know that p even though p has an epistemic probability of less than 1 for S. Our conception will be somewhat different. As we will understand it, to hold that fallible knowledge is possible is to hold that it is possible to know that p even though the method by which one formed one's belief that p could have led one to a false belief that p. Or, to put it in the framework of possible worlds: there is at least one possible world in which S

forms her belief that p by method M and p is false. That knowledge is fallible in this way is a very widely held view in epistemology. And with good reason; denying the possibility of fallible knowledge seems to lead to skepticism about the possibility of knowledge.[77]

The view that knowledge is simply safe belief will be understood as follows: S knows that p iff the method by which S formed her belief that p could not have easily led her to a false belief that p. In the possible worlds framework: there is no nearby possible world in which S falsely believes that p.[78] That safe belief is a necessary condition on knowing is a fairly popular view.[79] That safe belief is necessary and sufficient for knowledge is much more controversial.[80] It will nevertheless be taken as an assumption here.[81]

That belief is a functional kind is a dominant view in the philosophy of mind, and functionalists often take its functional role to be partly constituted by its connections to action. On this view it is commonly thought that S believes that p only if S is disposed to treat p as certain for the purposes of decision making in normal circumstances[82] [83], where to treat p as certain for decision making is to ignore ~p possibilities in one's decision making. Stronger

[77] Matters are made more complicated here if, as Williamson (2000) argues, one's evidence consists of all and only those propositions one knows to be true ('E=K'). If E=K is true, then the probability of p on one's evidence is always 1 when one knows that p (because p is part of one's evidence, and the probability of p on p is necessarily 1). If epistemic probabilities are understood as probabilities on one's evidence, then it follows that there is a sense in which knowledge is infallible. In that case much of what is argued in this section would need to be re-thought. E=K is not assumed here however (I argue against it in Hughes (2014)), and even were it to be accepted, further argument would be needed to motivate the thought that we should treat the kind of epistemic probability relevant to decision theory as probability on one's evidence.

[78] Some refinements will be needed here to deal with knowledge of necessary truths. Since there is no possible world in which necessary truths are false, on the current interpretation of the safety condition, anyone who believes a necessary truth also knows it. But this is too generous, for there are surely possible cases where S believes that p, and p is a necessary truth, but S does not know that p. For the relevant refinements, see Pritchard (2012). I will ignore this issue here for ease of exposition.

[79] Williamson (2000), Sosa (1999), and Pritchard (2005) all develop variations on this idea. It is not entirely uncontroversial though. For objections see Kelp (2009), Baumann (2008), Comesana (2005), and Neta & Rohrbaugh (2004).

[80] Williamson seems to be sympathetic to the view, and it is defended from objections by Lasonen-Aarnio (2010), but rejected by a number of epistemologists. For example, all those who reject the necessity direction in the last footnote, amongst others. Bonjour's (1980) clairvoyance cases are often thought to cause troubles for the sufficiency direction.

[81] It is worth noting that, given that it is a desiderata on a theory of knowledge that it is able to account for the value of knowledge, it may be thought that the account offered in this section provides us with a motivation for the safe belief view insofar as it can account for the value of knowledge. I won't push this line of thought here though.

[82] Ross & Schroeder (2014), Frankish (2004, 2009), Smithies (2012), Bratman (1987), and Wedgewood (2008) all endorse variations of this claim.

[83] But not necessarily *all* circumstances - for example, on the view we are interested in, one may count as believing that p even if one is not willing to take p as certain in a situation in which the stakes are very high and one has a great deal to lose if one is wrong about p. Obviously we will want to know more about what 'normal' circumstances are here. Unfortunately this issue is beyond the scope of the essay.

views are possible: that being disposed to take p as certain for decision making in normal circumstances is necessary *and sufficient* for believing p, or that being disposed to take p as certain for decision making in circumstances C is necessary for believing that p in C.[84] We only assume the weaker view here.

Ross and Schroeder (2014) explain why we should expect that, due to their cognitive and epistemic limitations, people have a need to employ a heuristic of treating propositions that are uncertain as certain for the purposes of decision making:

> "The need to treat uncertain propositions as true has long been acknowledged by decision theorists. Leonard Savage makes this point very clearly. According to his formulation of decision theory, decision problems are defined by a table in which 'consequences' (i.e., outcomes) are assigned to pairs consisting of acts and states of nature. Thus, the agent can calculate the *expected utility* of her possible acts (about whose actual utility she is uncertain) in terms of the *actual utilities* of the possible consequences of these acts, multiplied by the probabilities of the states of nature in which these acts would have these consequences. The general problem is that for any given act, an agent will typically have nonzero credence in vastly many possible consequences of this act. And so if she were to associate a given consequence with a given act-state pair only if she were certain that the act-state pair would have this consequence, then she would need to employ a vast partition of ultrafine-grained states of nature, and the resulting computational task would be unmanageable. Thus, Savage concludes that, inevitably, act-state pairs with "actually uncertain consequences [must] play the role of sure consequences" Savage [1972, p. 84]. Similar considerations apply, as James Joyce [1999] has shown, on other formulations of decision theory" (2014: 266)

Cast in simpler terms, the point is this. For almost every proposition one should, and typically will, have a non-zero credence that it is true, since there is almost nothing that we can be *completely* certain is false. If we were to factor all of these non-zero probabilities into our decision making, then coming to a decision about how to act would very often be an unmanageably complex task for creatures, like ourselves, who have limited cognitive

[84] Weatherson 2012 seems to hold this latter view.

capacities and often need to engage in decision making in time-pressured circumstances. Thus we have a need to simplify our decision making procedures in various ways. One way in which we do so is by sometimes treating uncertain propositions as certain, thereby simplifying decision making, and lightening the cognitive load, by eliminating the need to factor these uncertainties into the decision making task. Call this phenomenon of treating uncertainties as certain the 'uncertainty-certainty heuristic'. A moment's reflection will confirm that we do indeed employ this heuristic, at least in our conscious practical reasoning.

§2.4. The Argument

My suggestion is that if we put the view that knowledge is fallible safe belief together with the functionalist theory of belief and the uncertainty-certainty heuristic, we are availed of a plausible explanation for why knowledge is the norm of belief.

We begin with some entailments. First off, notice that if belief has the kind of constitutive connection to action suggested, then:

(a) It is permissible to believe that p iff one knows that p

entails:

(b) It is permissible to take p as certain in decision making in normal circumstances iff one knows that p[85]

And if knowledge is fallible safe belief, then (b) entails:

[85] This principle is similar to a number of 'knowledge-action principles' that have been proposed recently. For example, Hawthorne and Stanley (2008) argue that "Where one's choice is p-dependent, it is appropriate to treat the proposition p as a reason for acting iff you know that p". (b) differs from most of these proposals by not claiming that knowing that p is *always* sufficient for being in a strong enough epistemic position to treat p as certain in decision making. For example, it is compatible with (b) that when the stakes are very high one needs a stronger epistemic relationship to p than knowledge to permissibly treat p as certain in one's decision making. Arguably this is an advantage of (b) over Hawthorne and Stanley's proposal. As Brown (2008) and Reed (2010) have pointed out, where there is a great deal at stake about whether one is right about p, intuitively knowing that p is not good enough for one to permissibly treat p as certain. (b) does not face this problem. Knowledge-action principles like the one endorsed by Hawthorne and Stanley have been used to argue that knowledge is pragmatically encroached upon. No such argument can be run using (b).

(c) It is permissible to take p as certain in decision making in normal circumstances iff one could not have easily have falsely believed that p.

If we employ the uncertainty-certainty heuristic, it follows that:

(d) It is permissible to take an uncertainty as certain for the purposes of decision making in normal circumstances iff one could not easily have falsely believed that p.

I think there is a good explanation available for why (d) is true, and that this explains - at least in part - why it is permissible to believe that p iff you know that p.

The argument, enthymematically expressed, is this:

(1) Taking uncertainties as certain in decision making introduces epistemic risk to the actions that result from the decision procedure.

(2) Norms of decision making that allow for too much epistemic risk in action are unhelpful

(3) If (1) and (2), then we should expect to have a risk-reducing rule restricting when it is permissible to take uncertainties as certain in decision making

(4) Three desiderata on such a rule are:

a. It must reduce risk
b. It must not be too demanding so as to be unattainable
c. It must be cognitively cheap

(5) The knowledge rule fulfills these three desiderata very well.

So why is our epistemic system one in which it permissible to believe that p iff you know that p? Because a policy of believing that p iff you know that p is one that does a good job of enabling us to simplify decision making in a cognitively cheap way whilst preventing

excessively risky actions resulting from the simplification.

The claims of this argument will be argued for in turn. Beginning with (1), it should be obvious that treating uncertainties as certain in decision making results in riskier actions than only treating certainties as certain. The following example demonstrates this.

Suppose that S is playing roulette on a table that is rigged so that the ball is guaranteed to land red, and S is deciding whether to bet on red or black. Now suppose that S has super-reliable inside information that the table is rigged, and so it is epistemically certain for S that the ball will land red on the next spin and S knows this. Suppose that S reasons as follows:

(1) The ball will land red
(2) So, if I bet red I'll double my money, and if I bet black I'll lose my money
(3) I would prefer to double my money than lose my money
(4) So, I should bet red

And suppose that she proceeds to bet red.

Compare S to S*. Like S, S* is playing roulette on a table that is rigged so that the ball is guaranteed to land red, and is deciding whether to bet on red or black. Unlike S, S* as no evidence suggesting that the table is rigged. As such, it is just under .5 epistemically probable for her that the ball will land red. Suppose that S* reasons as follows:

(1) The ball will land red
(2) So, If I bet red I'll double my money, and if I bet black I'll lose my money
(3) I would prefer to double my money than lose my money
(4) So, I should bet red

Ignore the fact that S*'s reasoning is defective in that she should not be relying on (1), and focus on the following question: whose action is more risky? Clearly the answer is S*. And that's true *even though the ball will in fact land red in both cases*. The risk that S* runs is *epistemic*. Why is S*'s action more risky? Because there is a significantly higher chance, given her epistemic situation, that she will lose her money than there is for S. But S and S* reached their decisions in exactly the same manner. The only difference between S and S* is

that, unlike S, S* took an uncertainty as a certainty. Thus, the case demonstrates that the riskiness of acting on the basis of a decision making procedure is greater when one takes an uncertainty as certain than when one only take a certainty as certain.[86]

(2) claims that norms of decision making that allow for too much epistemic risk are unhelpful. Again, this should be obvious. Reasoning about how to act, rather than simply acting, is useful because it allows us to figure out which courses of action are better and which worse, given our preferences and information. The norms of decision making exist to guide you to better courses of action, and it is in virtue of their ability to do this that they have their normative force. If the norms of decision making were very lax when it came to the tolerance of epistemic risk, then conforming to them would often result in reasoners performing highly risky actions. Given the nature of high risk actions, this would often result in unwanted outcomes. Thus, were the norms of decision making to be very lax when it came to epistemic risk, they would fail to live up to their *raison d'etre* of guiding us towards better courses of action and away from worse ones. So we should expect them not to be too lax.

Why think that if (1) and (2) are true, then (3) will also be true? Because if taking uncertainties as certain introduces epistemic risk to the actions that result from decision making, and allowing for too much risk results in the norms of decision making failing to be a guide to the best course of action, then if the norms of decision making are to serve their purpose of guiding decision makers towards better courses of action, we should expect there to be rules governing the conditions under which it is permissible to take an uncertainty as certain in decision making that place restrictions on the level of epistemic risk that it is acceptable to take on in taking an uncertainty as certain. If there were no such rules, then the norms of decision making would not serve their purpose.

[86] Of course, it is also possible to end up doing something stupidly risky, not as a result of having a poor epistemic position with respect to the propositions you rely on in decision making, but rather because of your preferences. Imagine S**, who is in the same situation as S*, but reasons as follows:

(1) It is just under .5 likely that the ball will land red
(2) So there is a significant risk that if I bet all my money on red I'll lose all my money
(3) I want to run a significant risk of losing all my money
(4) So, I should bet all my money on red.

S** runs just as great a risk as S*, but the source of her risk is different. Her problem is not with the epistemic position that she bears to the propositions that she relies on in her decision making (we can suppose that each of them is epistemically certain for her and she knows this), but rather with her risk-loving preferences.

So we should expect there to be a risk-reducing rule (or set of rules) on when it is permissible to take uncertainties as certain in decision making. There will be at least three desiderata on a risk-reducing rule for it to be well suited to its task. Firstly, and most obviously, it must be one where the conditions to be met for permissibly taking an uncertainty as certain actually *do* reduce risk. Otherwise the rule will clearly not be fit for purpose.

Secondly, the conditions set down by the rule must be conditions that we human decision makers, with our limited epistemic abilities, can regularly meet. That is to say: they must not be too demanding. A rule that demanded that one take p as certain in decision making iff p is no more than 1/1 trillion epistemically likely to be false would certainly avoid the pitfall of allowing too much risk into decision making, but it would be useless on the plausible supposition that there is very little that is less than 1/1 trillion epistemically likely to be false for us, for it would be a rule that would be ill-suited to meet our needs. It would say, *'yes',* there are conditions under which you can simplify your decision making by treating an uncertainty as certain, but unfortunately for you these are conditions that almost never obtain. The result would be that we would only very rarely be allowed to simplify decision making in the relevant way, and hence that the uncertainty-certainty heuristic would be next to useless at serving its purpose of making otherwise unmanageably complex decision making tasks tractable for creatures, like ourselves, who have limited cognitive and epistemic capacities and often need to make decisions in time-pressured circumstances with limited information.

The third desideratum is that the rule must be cognitively cheap. By this I mean that the conditions laid down in the rule must be such that it is not a cognitively laborious task to establish whether they have been met. The reason for this desideratum is that the purpose of the risk-reducing rule is to allow decision makers to turn what would otherwise be a cognitively intractable decision making task into a manageable one by allowing them to treat uncertainties as certain whilst not taking on board too much risk. A rule that met the risk-reducing and achievability criteria would nevertheless be ill-suited to this role if it turned out to be a laborious cognitive task to establish whether the conditions laid down by the rule had been met on any given occasion, since this would vitiate the purpose of having a rule that *simplifies* decision making tasks in such a way that they can be undertaken in time-pressured circumstances with limited cognitive resources.

If it is permissible to believe that p iff you know that p, and belief has the functional role that

we are assuming it to have, then the risk-reducing rule is this: it is permissible to take an uncertainty as certain for the purposes of decision making in normal circumstances iff you know that p. That is to say: the condition laid down by the risk-reducing rule is that one must know that p. That the condition is knowledge, rather than something else, is, I want to argue, no surprise, since if knowledge is fallible safe belief, then it appears to fulfil our desiderata on a risk-reducing rule very well. And this is what explains (at least in part) why knowledge, rather than something else, is the norm of belief.[87]

Here's why the knowledge rule fulfils the desiderata well. First of all, notice that the first two desiderata will interact with one another. The need to avoid taking on excessive epistemic risk in the actions that result from decision making will put pressure on us to have a risk-reducing rule that requires one to be in a fairly good risk-avoiding epistemic position with respect to p in order to permissibly treat it as certain in decision making. That is to say, the need to avoid epistemic risk will push the required epistemic standards to be met *upwards*. But the need to have a rule whose conditions can be met by creatures such as ourselves will push the required epistemic standards *downwards*. If the standard to be met is *too* high, then it will be a standard that we can only rarely - if ever - meet, and as we have seen we need a standard that can be met by normal human epistemic agents on a regular basis. Given this upwards-downwards push and pull, we should expect the risk-reducing rule to be one that requires us to be in a strong, but nevertheless easily achievable, epistemic position to satisfy. If knowledge is fallible safe belief, then the knowledge rule of risk-reduction suits these demands very well. Knowledge in the image of fallible safe belief does not require that one could not have *possibly* falsely believed that p, but does require that one could not have *easily* falsely believed that p. This is a standard that we can, and do, achieve on a regular basis in our everyday lives. But it is also a standard that brings with it a good deal of risk-reduction. When one acts on a belief that could not have easily been false, one acts with a 'safety buffer' against the risk of a bad outcome. By contrast, when one acts on a belief that could have easily been false, one runs an uncomfortable level of risk, even if one's belief is both true and well supported by the evidence. To see this different, we can contrast the level of risk run by a subject who knows that p with the level of risk run by a subject who merely has a Gettiered rational true belief that p.

[87] I do not mean to suggest that *no* other rule could fulfill the desiderata. The questions of what other rules could fulfill them and why these aren't the norm of belief are beyond the scope of this essay. They are questions for further work.

Our example is borrowed from Brown (2008). Suppose, in the knowledge case, that S has arranged to meet her partner for lunch. S needs to take the train to meet him, and in the course of planning what time to leave the office, she consults the timetable on the train company's website. The timetable says there is a 12.15 express train that will get her to lunch on time. S thereby comes to know that there is a 12.15 express, and, on the basis of information, leaves the office shortly before 12.15, catches the train, and meets her partner on time.

In this case it would be rather odd to suggest that S ran a substantial risk of missing her lunch date in this case - what would the source of the risk be? Contrast this with a Gettier version of the situation. As before S has arranged to meet her partner for lunch. She needs to take the train to meet him, and in the course of planning what time to leave the office, she consults the timetable on the train company's website. The timetable says there is a 12.15 express train that will get her to lunch on time. On the basis of this information S leaves the office shortly before 12.15 in order to catch the train. Unbeknownst to S however, a hacker has accessed the train company's website and replaced this year's timetable with last year's timetable for a joke. Luckily for S, although the two timetables are very different, this year's timetable just happens to be the same as last year's when it comes to lunchtime express trains, so S makes the train and meets her partner on time.

Here, in contrast with the knowledge case, it is natural to think that in relying on an unreliable source of information, and thus a belief, which could have easily been false, that there is a 12.15 express train, S ran (albeit unwittingly) a substantial and undesirable risk of a bad outcome - namely missing the train and her lunch date. What this contrast between the knowledge case and the Gettier case demonstrates is that a safety condition on knowledge has the result that conforming to the knowledge rule ensures that one does not run an excessive level of risk in the actions that result from one's decision making and transgressing it involves placing oneself at substantial risk.

If knowledge is fallible safe belief then it does a good job of meeting the first two desiderata on a risk-reducing rule. It also does well on the third desiderata. It is a striking fact about knowledge attribution that it is a fairly effortless task. Consider the difference in the experience of trying to calculate '652 x 77 = ' with that of judging whether a subject in a given situation knows that p or not. The former is a slow, effortful, consciously transparent,

cognitively burdensome, task. In the parlance of dual-process theory, it is a System 2 process. By contrast, the latter is a quick, effortless, intuitive, and cognitively cheap, task that is not open to conscious inspection. The process of knowledge attribution has many of the hallmarks of a System 1 process. Accordingly, a knowledge rule of risk-reduction is well suited to meet the requirement of cognitive cheapness required by the third desiderata on a risk-reducing rule. The knowledge rule, unlike many other possible rules,[88] is one that can be applied in time-pressured circumstances by agents with limited cognitive resources.

That the knowledge rule on risk-reduction does a good job of fulfilling these three desiderata provides us, I want to suggest, with a good explanation for why knowledge is the norm of belief. To see this it will be helpful to recap the argument just given. We saw that if belief is a functional kind with the constitutive connections to decision making that we are assuming it has, then part of what it is to believe that p is to be disposed to treat p as certain for the purposes on decision making in normal circumstances. Thus, if it is permissible to believe that p iff you know that p, then it is permissible to treat p as certain for the purposes of decision making in normal circumstances iff you know that p. Given that people need to employ the uncertainty-certainty heuristic in order to make decision making procedures cognitively tractable, it follows that you are permitted to treat an uncertainty as certain for the purposes of decision making in normal circumstances iff you know that p. Treating uncertainties as certain brings with it epistemic risk, and we will have a need to avoid taking on too much risk. As a result, we will have a risk-reducing rule that fulfils the three desiderata of reducing risk, being achievable, and being cognitively cheap. The knowledge rule fulfils these three criteria very well, and that is why it should come as no surprise to us that the epistemic norm of belief is *knowledge* rather than something else. Or to put it more succinctly: part of what it is to believe is to often treat propositions as certain when they are uncertain. A good policy of taking uncertainties as certain will fulfil certain desiderata. The knowledge rule fulfils these desiderata very well, and this explains why it is permissible to believe that p iff you know that p. More succinctly still: a belief is fit to be relied on in decision making when, and only when, it is an item of knowledge.

[88] Consider, for example, a rule saying: treat p as certain iff p is at least *n* epistemically probable for you (where $n<1$). It's not easy to know how epistemically probable a proposition is for you. Establishing it will often involve marshalling a great deal of evidence, and even then you will often only be in a position to give a rough estimate. Accordingly, this rule would fail to meet the third desiderata and would thus not be a plausible candidate to be the risk-reducing rule.

As I said earlier, no pretence is made here that the argument of this section is conclusive. It relies on a number of assumptions, some of which - in particular the claim that knowledge is simply safe belief - are controversial. And it will stand or fall with those assumptions. Nor is any pretence made that this account *exhaustively* explains why knowledge is the norm of belief. It is perfectly compatible with everything that I have said that there are further reasons - perhaps more important or fundamental reasons - why our epistemic system is one in which it is permissible to believe that p iff one knows that p. Nevertheless, I suggest that the account sketched here represents a promising starting point in the project of answering the 'why?' question.

§III. Objections and Replies

We have covered a lot of ground, and there are many points in my arguments that one may object to. It is, of course, not possible to reply to every possible objection. Here I reply only to what I take to be the most pressing objections. Due to limitations of space, the focus will be on objections to the argument of §1 - that knowledge is the norm of belief - since this is the main claim that I want to defend.

Objection: You claim that were Anne to know that she didn't know whether there was cyanide or sugar in the jar, then she could not non-culpably spoon the substance into Bill's coffee without an excuse. And you claim that were Tom to know that he didn't know whether the man he has trained his gun on was a genuine armed robber or in fact innocent, then he could not non-culpably shoot the man without an excuse. The claim in these cases is that Anne and Tom can only permissibly Φ if they *know* that Φ-ing is the objectively right thing to do. Whilst this has some *prima facie* plausibility, it overlooks the fact that Anne and Tom could find themselves in situations in which, whilst they don't know that Φ-ing is the objectively

right thing to do, it is overwhelmingly likely given their evidence that Φ-ing is the objectively right thing to do. And in such circumstances it is plausible that they would be permitted to Φ, despite not knowing that Φ-ing is the objectively right thing to do. Accordingly, the cases of Anne and Tom don't support your argument.

Reply: I rather doubt that Anne and Tom can non-culpably Φ without an excuse solely on the basis of the probabilities involved, without knowing that Φ-ing is the objectively right thing to do. But if you are unhappy with these cases, the following adaptation of the *Cyanide Coffee* case ought to bring out the point that there are some actions which one can only permissibly perform when one knows that they are objectively right. Suppose that there are two jars side by side in Anne's kitchen, and that, as before, she wants to spoon sugar into Bill's coffee. Suppose that Anne knows that jar A contains sugar. Suppose that Anne does not know that jar B contains sugar rather than cyanide, but that it is highly probable on her evidence that it contains sugar; a lottery was run by Anne's housemate to determine whether to fill jar B with sugar or cyanide, and as the lottery was set up there was .99999999 chance that the jar would be filled with sugar, and a .00000001 chance that it would be filled with cyanide. Anne knows about the lottery, but doesn't know what the result was. Nevertheless, she knows that it is .99999999 likely on her evidence that jar B contains sugar. Which jar should Anne choose to fill Bill's coffee from? Clearly the answer is jar A. And were she instead to choose jar B, we would surely hold Anne culpable for needlessly putting Bill at risk (unless she had an excuse). Anne is not permitted to choose jar B. Notice however, that had Anne *known* that jar B contained sugar, she would have be permitted to choose jar B, and could have non-culpably done so without an excuse. In Anne's situation, where she has the option of choosing between a course of action that she knows is objectively right (selecting jar A) and a course of action that she does not know is objectively right, despite the high probability of it being so (selecting jar B), it is not permissible for her to take the course of action that she does not know to be objectively right. So if you are unhappy with the original Anne and Tom cases, simply replace them with this case.

Objection: Here's a reason to think that the 1st permission principle is false. Some acts are permissible, but ill advised. When someone performs a permissible but ill advised act, they are often properly criticisable. For example, it may well be that you are permitted, as an adult, to eat junk food all the time, if that's what you want to do. But you shouldn't expect not to be criticised by your family for doing so. Given that permissible but ill advised acts have the

property of being properly criticisable, it can be argued that, contrary to your claims, believing without knowing is permissible, since it can be argued that believing without knowing is permissible, but ill advised, and *this* explains why you are criticisable when you believe that p whilst knowing that you don't know that p. Why shouldn't we think this is what is going on?

Reply: We should draw a distinction between being permitted-but-criticisable and being culpable-because-not-permitted-without-an-excuse. Believing 'p, but I don't know that p' belongs in the latter camp. Not believing 'p, but I don't know that p'' is a *requirement* of rationality, not a recommendation. When Φ-ing is merely recommended, rather than required we typically are in a position to concede that, 'yes, it's up to you, even if I think it's a bad idea.' Not so with 'p, but I don't know that p'. If you believe 'p, but I don't know that p' you've fallen into irrationality of a kind that it not merely ill advised, but straightforwardly prohibited, from an epistemic perspective. When we say to someone who has just acknowledged that they don't know that p - 'well, you shouldn't outright believe that p then' - the force of the 'shouldn't' here is not one of recommendation, but rather prohibition.

Objection: You've argued that ignorance can excuse impermissible acts, and so render the person who performs them non-culpable. That's true, but there is a flipside to it that you haven't mentioned, and it causes problems for your argument. Just as there can be blameless transgressions, so to can there be *reckless adherences* - that is, situations where Φ-ing is permissible, but your evidence suggests that it isn't, where you nevertheless go ahead and Φ. For example, suppose that the speed limit on the road is 70mph, and you are doing 65mph. In that case, you are driving at a permissible speed. But all your evidence suggests that you are actually doing 100mph. In that case you recklessly adhere to the speed limit, and intuitively you are culpable unless you have an excuse. The same goes for believing. If you are in fact permitted to believe that p, because you know that p, but all of your evidence suggests that you don't know that p, then you are culpable for your risk-taking *qua* epistemic agent. If so, then the 1st permissibility principle is false - for here we have a case where Φ-ing is permissible, but you cannot non-culpably Φ without an excuse. Once we see the possibility of such states-of-affairs, your argument collapses, because you either have to deny the permission principles that your argument appeal to - in which case the knowledge norm is unmotivated - or you keep the principles, but are then forced to infer that one is permitted to believe only if you know that you know that p. But that's too strong. Knowledge is sufficient

for permissible belief. So if you keep the principles, your argument overgenerates. Either way, your argument fails.

Reply: I think that believing when you know but it is unlikely that you know isn't culpable risk-taking. Anti-risk conditions are *built in* to the permissibility conditions for moral and epistemic norms. The speeding case is misleading. In the speeding case, you are *legally* permitted to do what you do, and not *legally* culpable for your behaviour. But that doesn't mean you're not *morally* culpable for it. Morally speaking, sticking within the speed limit is not sufficient for permissible driving (even if it is sufficient for legally permissible driving). You are not permitted to recklessly adhere to the speed limit. The same goes for other such cases. And similarly for believing. The anti-risk condition is *built in* to the permissibility conditions for belief. As we saw in §2, it is plausible that it is for this very reason that you have to *know* that p, rather than merely have a true belief in order to permissibly believe that p. But it is not required that you be *safely* non-risky in your believing (or, as we may put it, 'non-riskily non-risky'). And as such, you are not culpable for believing that p in the case where you know that p, but your evidence suggests that you don't know that p. The fact that *you could easily* have acted riskily is not relevant to the assessment of whether you are blameworthy. Why, then, does it appear as though you *are* culpable for believing in such improbable knowledge cases? One suggestion, to which I am sympathetic, is that the possibility of improbable knowledge is highly counterintuitive, and so we tend to treat cases of improbable knowledge as though they were cases of non-knowledge (Williamson (2009)), and so as cases of culpably risky believing. So I see no problem in holding on to both the *1st permissibility principle* and the claim that knowledge is sufficient for permissible belief.

Objection: According to the *3rd permissibility principle* Φ-ing without knowing that Φ-ing is objectively right is permissible for S iff were S to know that she did not know that Φ-ing is objectively right, it would nevertheless be possible for S to non-culpably Φ without an excuse. On the standard Lewis-Stalnaker analysis of subjunctive conditionals they are true whenever the antecedent is impossible. It follows that whenever it is impossible to know that one does not know that Φ-ing is objectively right, one is permitted to perform every possible Φ without knowing that Φ-ing is objectively right. That would be highly implausible, since there clearly are actions that one is not permitted to perform unless one knows that they are objectively right. Moreover, it would deliver the result that it is permissible to believe that p without knowing that p, which is exactly what you want to argue against. The problem for you is that

there *are* situations in which it is impossible to know that you don't know whether Φ-ing is objectively right. Two such circumstances are: 1. borderline cases of knowing whether Φ-ing is objectively right, and 2. cases where you rationally believe, without knowing, that Φ-ing is objectively right. The upshot of this is that the *3rd permissibility principle* has absurdly implausible consequences about what is and isn't permissible. So it, and the argument for the knowledge norm of belief that rests on it, should be rejected.

Reply: I accept the Lewis-Stalnaker analysis of subjunctive conditionals, but I deny that this causes the kind of problem for my view that the objection claims. With respect to the point about borderline cases of knowledge being unknowable, it is crucial to recognise that, as Srinivasan (forthcoming) has emphasised, luminosity failure is a contingent feature of a subject that is the result of the limited discriminatory capacities of that subject. There could, in principle, be a subject with perfect discrimination, and so for whom every condition, including knowledge, is luminous. In assessing the subjunctive conditional: 'were S to know that she did not know that Φ-ing is objectively right she could have nevertheless non-culpably Φ'ed without an excuse', nothing forces us to keep fixed S's discriminatory capacities. Thus, we can imagine an *idealised* version of S when assessing the conditional. The thought is that whilst, for normal-Bob, who has average human discriminatory capacities, there may be situations in which Φ-ing is permissible despite the fact that Bob does not know that Φ-ing is objectively right, but Bob is not in a position to know that he does not know that Φ-ing is objectively right since he is unable to discriminate between the case where he doesn't know and a similar case where he does know, the same is not true of idealised-Bob, who has perfect discriminatory capacities, since idealised-Bob *can* discriminate between the situation where he does not know that Φ-ing is objectively right and a similar case where he does know. Thus we can sensibly judge whether normal-Bob is permitted to Φ by asking whether idealised-Bob could non-culpably Φ without an excuse. As a result, I don't think that borderline cases cause the kind of problem for my view that the objection claims.

This response isn't available in the rational-believe-without-knowledge case; if rationally believing that p is incompatible with knowing that you don't know that p, then no matter how good your discriminatory capacities are, it will not be possible to know that you rationally believe that p and know that you don't know that p. Nevertheless, rational-believe-without-knowledge cases also fail to imperil the view. The objection is that it is impossible to know that you don't know that p when you rationally believe that p, and as such the subjunctive

conditional will be necessarily true in this circumstance. However, whilst it is true that it is impossible for you to know that you don't know that p whilst you rationally believe that p, it is only impossible *when it is held fixed that you rationally believe that p.* But nothing requires us to hold this aspect of your circumstances fixed in assessing the conditional, and it is of course possible to acquire new evidence that puts you in a position to know that you don't know that p. Of course, once you have acquired this evidence it will no longer be rational for you to believe that p, but this is exactly what we should expect if the *3rd permissibility principle* is true and it is irrational to believe that p whilst knowing that you don't know that p.

Objection*:* Your argument departs from the assumption that it is irrational to believe that p if you know that you don't know that p, and that, it is claimed, derives from the fact that it is irrational to believe 'p, and I don't know that'. But the former follows from the latter only on the assumption that rational belief is closed under conjunction. What is irrational is believing the *conjunction* p, but I don't know that p', not believing p and also believing that you don't know that p. And rational belief isn't closed under conjunction. Just because S rationally believes that p, and rationally believes that q, it doesn't follow that S rationally believes p&q (or even believes p&q). So the assumption that the argument for the knowledge norm of belief departs from - that it is irrational to believe that p if you know that you don't know that p, does not follow, as you claim, from the irrationality of believing 'p, but I don't know that p'. So your argument stands in need of further support.

Reply: There are two main arguments for rejecting the claim that rational belief is closed under conjunction - one stemming from a Lockean theory of belief, and the other from the preface paradox. Neither, I think, causes trouble for my argument. One motivation for thinking that rational belief isn't closed under conjunction stems from a Lockean theory of belief whereby S believes that p iff S has a credence in p above a threshold that stands somewhere short of 1, since the joint probability of p&q is less than the individual probabilities of p and q if the probability of p is <1 and the probability of q is <1. I reject this view of belief for a number of reasons, which I won't go into here. More importantly however, it is anyway anathema to my starting assumption - that it is irrational to believe 'p but I don't know that p' - since it can surely be rational to have a arbitrarily high credence in p whilst knowing that you don't know that p - I take it that lottery cases show just this. So if you think rational belief isn't closed under conjunction, then you should think that the inference from the irrationality of believing p, but I don't know that p' to the irrationality of believing that p

when you know that you don't know that p, is the least of my worries - you should reject my starting assumption that it is irrational to believe 'p, but I don't know that p', since it is, I take it, possible to know that you don't know that your lottery ticket has lost, prior to the draw, even when the probability that it has lost is arbitrarily high short of 1.

Another motivation for denying the closure of rational belief under conjunction comes from the Preface Paradox: it seems plausible that an author could rationally believe all of the assertions in her book, yet also rationally believe that the book contains at least one false assertion. If so, then rational belief isn't closed under conjunction. However, even if this is correct, notice that the preface paradox derives its plausibility from the fact that the conjunction is *very long*, and so likely to contain at least one error. But the irrationality of believing 'p, but I don't know that p' involves only two conjuncts, so the kinds of failures of closure of rational belief under conjunction that the preface paradox motivates do not concern my argument.

So I don't think that any of the main motivations for denying that rational belief is closed under conjunction trouble my argument.

Objection: just as it is irrational to believe 'p, and I don't know that p', so it is irrational to believe:

- p, and I can't be certain that p
- p, and I can't be sure that p
- p, but it is possible that not p

And we could run arguments parallel to the one you gave for the knowledge norm of belief, using these conjunctions, to get the conclusion that you are permitted to believe that p only if you can be certain that p/can be sure that p/it is not possible for you that not p. But that's far too strong. Knowledge is fallible, and knowledge is sufficient for permissible belief. So the kind of argument you give overgenerates, since, if it is accepted, it motivates demands on permissible belief that are too strong.

Reply: As Williamson (2000) points out, people are generally reluctant to let the contextually set standards for knowing come apart from those of being epistemically certain, having a right

to be sure, and it's being possible that not p. 'He knows that p, but he can't be certain', 'He knows that p, but he isn't sure that p', and 'He knows that p, but it's possible for him that not p' all sound bad. We are happy to attribute epistemic certainty, the right to be sure, and the impossibility of being wrong when *strictly speaking* they are not present. We do so when the subject knows. This is not surprising, given the notion of outright belief that knowledge is normatively necessary and sufficient for is one that constitutes a kind of contextual certainty.

Objection: Even if your response to the stronger Moorean conjunctions objection - that people don't let the contextually set standards for knowledge and certainty come apart - is accepted, there is another group of Moorean sounding conjunctions that cause problems for you. Namely:

- p, but I don't know whether I know that p
- p, but I don't know that I know that p
- p, but it's possible I don't know that p
- p, but it's unlikely that I know that p

The worry is this. These sound Moorean in much the same way that 'p, but I don't know that p' sounds Moorean. If so, then oughtn't we conclude that it is irrational to believe them? And if it is, then we can run an argument parallel to your argument to the conclusion that it is permissible to believe that p only if you know that you know that p. But again, that's too strong. So your argument overgenerates.

Reply: My response here is flat-footed. It is rational to believe what you know, even if it is unlikely that you know it. So I do not think that it is irrational to believe these conjunctions. Rationality, just like any other condition, isn't luminous. That's something we have to learn to live with. 'p, but it's unlikely that it's rational for me to believe that p' sounds just as bad, but we have to accept that it's possible to rationally believe even when it is unlikely that it is rational to believe, otherwise we commit ourselves to a tolerance principle, and with it, scepticism about the possibility of rational belief.

Objection: You are still not out of the woods with Moorean conjunctions. One can know that p even though it is very improbable on one's evidence that one knows that p. Similarly, one can rationally believe that p, even though it is very improbable on one's evidence that one

rationally believes that p. Given this it can be rational to have a high credence in the following conjunctions:

- p, and my evidence doesn't support that p
- p, and my evidence supports ~p
- p, and it is not rational for me to believe that p

- But these are just as bad - arguably worse - than 'p, and I don't know that p'. So why should we accept your starting assumption that it is irrational to outright believe 'p, but I don't know that p', if it can be rational to believe these kinds of Moorean conjunctions?

Reply: My starting assumption was that it is irrational to *outright believe* Moorean conjunctions, not that it is irrational to have a high credence in them. But as we saw in §2.5 of essay 4 ('Excuses and Epistemic Norms'), the primary argument for the possibility of improbable rational belief does not support the conclusion that it can be rational to outright believe that one does not rationally believe that p. Accordingly there is no present reason to think that one could rationally outright believe these conjunctions. As such, I do not think they cause problems for my view.

§IV. Conclusion

It is irrational to believe 'p, but I don't know that p'. This fact, I have argued, implies that knowledge is the norm of belief in the sense that it is permissible to believe that p only if you know that p. Why? Because if it is irrational to believe 'p, but I don't know that p', then it is only possible to non-culpably to believe that p without knowing that p when you have the excuse of not knowing that you don't know that p, and were it permissible to believe that p without knowing that p, then you would not need an excuse in order to non-culpably do so.

Since it is uncontroversial that it is permissible to believe that p if you know that p, this can be strengthened to the claim that knowledge is the norm of belief in the sense that it is permissible to believe that p *iff* you know that p. Why is knowledge the norm of belief in this sense? One reason, I have suggested, is because a belief is fit to be relied on in decision making when, and only when, it is an item of knowledge.

6. References

Adler, J. 2002, *Belief's Own Ethics* (MIT Press)

Alexander. J, Gonnerman. C, & Waterman. J, *forthcoming*, "Salience and Epistemic Egocentrism" in Bebee. J, ed *Advances in Experimental Epistemology* (Continuum)

Bach, K. 1985, "A Rationale for Reliabilism" in *The Monist* Vol. 68 (2). pp. 246-263

Bach, K. 2010, "Knowledge in and out of context" in Campbell. J, O'Rourke. M, & Silverstein. H, eds, *Knowledge and Skepticism* (MIT Press)

Baron, J. & Hershey, J. 1988, "Outcome Bias in Decision Evaluation" in *Journal of Personality and Social Psychology* Vol. 54. pp. 569-579

Baumann, P. 2008, "Is Knowledge Safe?" in *American Philosophical Quarterly* Vol. 45 (1). pp. 19-30

Birch, S.A. *et al,* 2004, "Understanding children's and adult's limitations in mental state reasoning" in *Trends in Cognitive Science* Vol. 8 (6). pp. 255-60

Bird, A. 2007, "Justified judging" in *Philosophy and Phenomenological Research* Vol. 74 (1). pp. 81-110

Bonjour, L. 1980, "Externalist Theories of Empirical Knowledge" in *Midwest Studies in Philosophy* Vol. 5 (1). pp. 53-74

Bratman, M. 1987, *Intention, Plans, and Practical Reason* (Cambridge: Harvard University Press)

Brown, J. 2008, "Subject-Sensitive Invariantism and the Knowledge Norm of Practical Reasoning" in *Nous* Vol. 42. pp. 167-189

Brown, J. 2013, "Experimental Philosophy, Contextualism, and SSI" in *Philosophy and Phenomenological Research* 86 (2). pp. 233-261

Brueckner, A. 2005, "Knowledge, Evidence, and Skepticism According To Williamson" in *Philosophy and Phenomenological Research* Vol. 70 (2). pp. 436-443

Buckwalter, W. 2010, "Knowledge Isn't Closed on Saturday: A Study in Ordinary Language" in *The Review of Philosophy and Psychology* Vol. 1 (3). pp. 395-406

Byrne, R. *et al*, 1999, "Counterexamples and the suppression of inferences" in *Journal of Memory and Language* Vol. 40. pp. 347-373

Byrne, R. *et al*, 2000, "The temporality effect in counterfactual thinking about what might have been" in *Memory & Cognition* Vol. 28. pp. 264-281

Byrne, R. 2002, "Mental models and counterfactual thoughts about what might have been" in *Trends in Cognitive Sciences* Vol. 6 (10). pp. 426-431

Byrne, R. 2005, *The Rational Imagination* (MIT Press)

Camerer, C. *et al,* 1989, "The Curse of Knowledge in Economic Settings: An Experimental

Analysis" in *Journal of Political Economy* Vol. 97 (5). pp. 1232-1254

Chisholm, R. 1966, *Theory of Knowledge*. (Englewood Cliffs; NJ: Prentice Hall)

Christensen, D. 2010, "Higher Order Evidence" in *Philosophy and Phenomenological Research* Vol. 81 (1). pp. 185-215

Cohen, S. 1984, "Justification and Truth" in *Philosophical Studies* Vol. 46. pp. 279-295

Cohen, S. 2002, "Basic Knowledge and the Problem of Easy Knowledge" in *Philosophy and Phenomenological Research*. Vol. 65 (2). pp. 309-329

Cohen, S. manuscript. "Theorizing About the Epistemic"

Cohen, S. & Comesana, J. 2013a, "Williamson on Gettier Cases and Epistemic Logic" in *Inquiry: An Interdisciplinary Journal of Philosophy* Vol. 54 (1). pp. 15-29

Cohen, S. & Comesana, J. 2013b, "Williamson on Gettier Cases in Epistemic Logic and the Knowledge Norm for Rational Belief: A Reply to a Reply to a Reply" in *Inquiry: An Interdisciplinary Journal of Philosophy* Vol. 56 (4). pp. 400-415

Conee, E. & Feldman, R. 2004, *Evidentialism* (Oxford University Press)

Conee, E. & Feldman, R. 2008, "Evidence". In Q. Smith (*ed*.), *Epistemology: New essays* pp. 83–104. (New York: Oxford University Press)

Comesana, J. 2005, "Unsafe Knowledge" in *Synthese* Vol. 146. pp. 395-404

Comesana, J. & Kantin, H. 2010, "Is evidence knowledge?" in *Philosophy and Phenomenological Research*. Vol. 80. pp. 447–454.

Costello, F. & Keane, M. 2000, "Efficient creativity: Constraint guided conceptual combination" in *Cognitive Science*. Vol. 24 (2). pp. 299-349

Craig, E. 1990, *Knowledge and the State of Nature: An Essay in Conceptual Synthesis* (Oxford: Oxford Clarendon Press)

Dancy, J. 2011, "Acting In Ignorance" in *Frontiers of philosophy in China* Vol. 6 (3). pp. 345-357

de Almedia, C. 2001, "What Moore's Paradox Is About" in *Philosophy and Phenomenological Research*. Vol 62. pp. 33-58

DeRose, K. 1991, "Epistemic possibilities" in *Philosophical Review* Vol. 100. pp. 581-605

DeRose, K. 2002, "Assertion, Knowledge, and Context" in *Philosophical Review* Vol. 111. pp. 167-203

DeRose, K. 2005, "The Ordinary Language Basis for Contextualism and the New Invariantism" in *The Philosophical Quarterly* Vol. 55 (219). pp. 172-198

DeRose, K. 2009, *The Case for Contextualism* (Oxford: Oxford University Press)

Dodd, D. 2007, "Why Williamson Should Be A Sceptic" in *The Philosophical Quarterly* Vol.

57. pp. 635-649

Elga, A. 2013, "The puzzle of the unmarked clock and the new rational reflection principle" in *Philosophical Studies* Vol. 164 (1). pp. 127-139

Fantl, J. & McGrath, M. 2002, "Evidence, Pragmatics, and Justification" in *Philosophical Review* 111 (1). pp. 67-94

Fantl, J. & McGrath, M. 2007, "On Pragmatic Encroachment in Epistemology" in *Philosophy and Phenomenological Research* Vol. 75 (3). pp. 558-589

Fantl, J. & McGrath, M. 2009, *Knowledge in an Uncertain World* (Oxford University Press: Oxford)

Feltz, A. & Zarpentine, C. 2010, "Do You Know More When It Matters Less?" in *Philosophical Psychology*, Vol. 23 (5). pp. 683-706

Fink, R. *et al,* 1992, *Creative Cognition: Theory, Research, and Applications* (MIT Press)

Foley, R. 1993, *Working without a net* (Oxford: Oxford University Press)

Frankish, K. 2004, *Mind and Supermind* (Cambridge: Cambridge University Press)

Frankish, K. 2009, "Partial Belief and Flat Out Belief" in Huber, F & Schmidt-Petri, C (*eds.*) *Degrees of Belief* (Synthese Library). pp. 75-93.

Gardner, J. 1997, "The Gist of Excuses" in *Buffalo Criminal Law Review* Vol. 2 (1)

Gerken, M. 2011, "Warrant and Action" in *Synthese* Vol. 178 (3). pp. 529-547

Gerken, M. 2013, "Epistemic Focal Bias" in *Australasian Journal of Philosophy* Vol. 91 (1). pp. 41-61

Gibbons, J. 2013, *The Norm of Belief* (Oxford: Oxford University Press)

Gigerenzer, G. & Todd, P. 1999, *Simple Heuristics That Make Us Smart* (Oxford: Oxford University Press)

Goldman, A. 1976, "Discrimination and perceptual knowledge" in *Journal of Philosophy* Vol. 73. pp. 771-91

Goldman, A. 2009, "Williamson on knowledge and evidence" in P. Greenough & D. Pritchard (Eds.), *Williamson on knowledge* pp. 73–92. (New York: Oxford University Press)

Greco, D. 2014, "A puzzle about epistemic akrasia" in *Philosophical Studies* Vol. 167 (2). pp. 201-219

Harman, E. 2011, "Does Moral Ignorance Exculpate? in *Ratio* Vol. 24 (4). pp. 443-468

Harman, G. 1973, *Thought* (Princeton: Princeton University Press)

Hawthorne, J. 2004, *Knowledge and Lotteries* (Oxford: Oxford University Press)

Hawthorne, J. & Srinivasan, A. 2013, "Disagreement without Transparency: Some Bleak Thoughts" in Lackey, J. & Christensen, D. (eds.) *The Epistemology of Disagreement* (Oxford:

Oxford University Press)

Hawthorne, J. & Stanley, J. 2008, "Knowledge and action" in *Journal of Philosophy* Vol. 105 (10) pp. 571-590

Heal, J. 1994, "Moore's Paradox: A Wittgensteinian Approach" in *Mind* Vol. 103 (409). pp. 5-24

Hegarty, P. & Pratto, F. 2000, "The effects of social category norms and stereotypes on explanation for intergroup differences" in *Journal of Personality and Social Psychology*, Vol. 80. pp. 723-735

Hetherington, S. 1998, "Actually knowing" in *Philosophical Quarterly* Vol. 48. 453-69

Hill, C. & Schechter, J. 2007, "Hawthorne's Lottery Puzzle and the Nature of Belief" in *Philosophical Issues*. Vol. 17. pp. 102-122

Hintikka, J. 1962, *Knowlege and belief* (Ithaca: Cornell University Press)

Hornsby, J. 2007, "Knowledge in action" in Leist, A. ed. *Action in Context* (Berlin: De Gruyter)

Huber, F. & Schmidt-Petri, C. 2009, (*eds.*), *Degrees of Belief* (Synthese Library)

Huemer, M. 2007, "Moore's Paradox and the Norm of Belief" in Nuccetelli, S & Seay, G (*eds.*) *Themes from G.E Moore* pp. 142-158 (Oxford: Oxford University Press)

Hughes, N. 2014. "Consistency and Evidence" in *Philosophical Studies* Vol. 169 (2). pp. 333-338

Horowitz, S. 2013, "Epistemic Akrasia" in *Nous* Vol. 48 (3). pp. 718-744

Hyman, J. 1999, "How knowledge works" in *Philosophical Quarterly* Vol. 49 (197). pp. 433-491

Hyman, J. 2006, "Knowledge and Evidence" in *Mind* Vol. 115 (460). pp. 891-916

Ichikawa, J. 2012, "Knowledge Norms and Acting Well" in *Though: A Journal of Philosophy* Vol. 1 (1). pp. 49-55

Ichikawa, J. 2014, "Justification is Potential Knowledge" in *Canadian Journal of Philosophy* Vol. 44 (2). pp. 184-206

Johnson-Laird, P. & Byrne, R. 2002, "Conditionals: a theory of meaning, pragmatics, and inference" in *Psychology Review* Vol. 109 (4). pp. 646-678

Jones, O. R. 1991, "Moore's paradox, assertion, and knowledge" in *Analysis* Vol. 51. pp. 183-186

Joyce, J. 1999, *The Foundations Of Causal Decision Theory*, (Cambridge: Cambridge University Press)

Joyce, J. 2004, "Williamson on Evidence and Knowledge" in *Analytic Philosophy* Vol 45 (4).

pp. 296-305

Kahneman, D. & Miller, D. 1984, "Norm Theory" in *Psychological Review* Vol. 93 (2). pp. 136-153

Kahneman, D. & Tversky, A. 1972, "Subjective probability: A judgment of representativeness" in *Cognitive Psychology* Vol. 3. pp. 430-454

Kahneman, D. & Tversky, A. 1973, "Availability: A heuristic for judging frequency and probability" in *Cognitive Psychology* Vol. 5. pp. 207-232

Kahneman, D. & Tversky, A. 1982, "The simulation heuristic" in Kahneman ed., *Judgment Under Uncertainty: Heuristics and Biases* pp. 201-208 (Cambridge University Press)

Kelp, C. 2009, "Knowledge and Safety" in *Journal of Philosophical Research* Vol. 34. pp. 21-31

Klein, P. 1985, "The virtues of inconsistency" in *The Monist* Vol. 68. 105-135

Krueger, J. & Clement, R. 1994, "The truly false consensus effect: An ineradicable and egocentric bias in social perception" in *Journal of Personality and Social Psychology*. Vol. 67 (4). pp. 596-610

Kvanvig, J. 2010, "Norms of Assertion" in *Assertion* pp. 233-250 (*eds*.) Brown, J. & Cappelen, H. (Oxford University Press)

Kyburg, H. 1970, "Conjunctivitis" in Swain, M. ed. *Induction, Acceptance, and Rational Belief* (Dordrecht: Reidel)

Lackey, J. 2007, "Norms of Assertion" pp. 594-626 in *Nous* Vol. 41: 4

Lasonen-Aarnio, M. 2010, "Unreasonable Knowledge" pp. 1-21 in *Philosophical Perspectives* Vol. 24

Lasonen-Aarnio, M. 2014, "Higher-Order Evidence and the Limits of Defeat" pp. 314-345 in *Philosophy and Phenomenological Research* Vol. 88 (2)

Lewis, D. 1996, "Elusive knowledge" in *Australasian Journal of Philosophy* Vol. 74 (4). pp. 549-567

Lichtenstein, S. *et al*, 1978, "Judged frequency of lethal events" in *Journal of Experimental Psychology: Human Learning and Memory* Vol. 4 (6). pp. 551-578

Littlejohn, C. 2009, "Must we act only on what we know?" in *Journal of Philosophy* Vol. 106 (8). pp. 463-473

Littlejohn, C. 2010, "Moore's Paradox and Epistemic Norms" in *Australasian Journal of Philosophy* Vol. 88 (1). pp. 79-100

Littlejohn, C. 2012, "Evidence and Knowledge" in *Erkenntnis* Vol. 74. pp. 241-262.

Littlejohn, C. 2012. *Justification and the Truth-Connection* (Cambridge: Cambridge

University Press)

Littlejohn, C. 2013, "The Russellian Retreat" in *Proceedings of the Aristotelian Society* Vol. 113 (3). pp. 293-320

Lycan, W. 2006, "On the Gettier Problem problem" in *Epistemology Futures*, (*ed*.) S. Hetherington (Oxford: Oxford University Press)

MacFarlane, J. 2005, "Knowledge Laundering: Testimony and Sensitive Invariantism: in *Analysis* Vol. 65, pp. 132-138

Markman, K. & Tetlock, P. 2000, "Accountability and Close-Call Counterfactuals: The Loser Who Nearly Won and the Winner Who Nearly Lost" in *Personality and Social Psychology Bulletin* Vol. 28 (10). pp. 1213-1224

McGlynn, A. 2014, *Knowledge First?* (Palgrave McMillan)

Miller, D. & Gunasegaram, S. 1990, "Temporal order and the perceived mutability of events: implications for blame assignment" in *Journal of Personality and Social Psychology* Vol. 59. pp. 1111-1118

Moore, G. E. 1962, *Commonplace book: 1919-1953* (London: George Allen & Unwin, Ltd.)

Nagel, J. 2010**,** "Knowledge Ascriptions and the Psychological Consequences of Thinking about Error" in *Philosophical Quarterly* Vol. 60 (239). pp. 286-306

Neta, R. 2008. "What Evidence Do You Have?" in *British Journal for Philosophy of Science* Vol 59. pp. 89-119

Neta, R. 2009, "Treating something as a reason for action" in *Nous* Vol. 43. pp. 684-699

Neta, R. & Rohrbaugh, G. 2004, "Luminosity and the Safety of Knowledge" in *Pacific Philosophical Quarterly* Vol. 85. pp. 396-406

Nozick, R. 1981, *Philosophical Explanations*, (Harvard University Press)

Oppenheimer, D. *et al*, 2004, "Spontaneous Discounting of Availability in Frequency Judgment Tasks" in *Psychological Science* Vol. 15. pp. 100-105

Parfit, D. 2011, *On What Matters Vol.1* (Oxford: Oxford University Press)

Pinillos, A. 2012, "Knowledge, Experiments, and Practical Interests" in Gerken, M. & Brown, J. (*eds.*) *Knowledge Ascriptions* (Oxford: Oxford University Press)

Pohl, R. & Hell, W. 1996, "No Reduction in Hindsight Bias After Complete Information and Repeated Testing" in *Organizational Behaviour and Human Decision Processes* Vol. 67. pp. 49-58

Pritchard, D. 2005, *Epistemic Luck* (Oxford University Press)

Pritchard, D. 2007, "Sensitivity, safety, and anti-luck epistemology" in Greco, J. ed. *The Oxford Handbook of Scepticism* (Oxford: Oxford University Press)

Pritchard, D. 2008, "Radical scepticism, Epistemic Luck, and Epistemic Value" in *Proceedings of the Aristotelian Society (suppl. vol.)* 82. pp. 19-41

Pritchard, D. 2010, "Knowledge and Understanding" in *The Nature and Value of Knowledge: Three Investigations*, D. H. Pritchard, A, Millar. & A, Haddock. (*eds.*) (Oxford: Oxford University Press)

Pritchard, D. 2012, "Anti-Luck Virtue Epistemology" in *Journal of Philosophy* Vol. 109 (3). pp. 247-279

Read, D. 1985, *Determinants of relative mutability*, Unpublished research.

Reed, B. 2010, "A Defence of Stable Invariantism" in *Nous* 44 (2). pp. 224-244

Reynolds, S. 2013, "Justification as the appearance of knowledge" in *Philosophical Studies* Vol. 163 (2). pp. 367-383

Ross, J. & Schroeder, M. 2014, "Belief, Credence and Pragmatic Enchroacment" in *Philosophy and Phenomenological Research* Vol. 88 (2). pp. 259-288

Royzman, E. *et al*, 2003, ""I know, you know": Epistemic egocentrism in children and adults" in *Review of General Psychology* Vol. 7 (1). pp. 38-65

Savage. L, 1972, *The Foundations of Statistics* (Dover Publishing)

Schaffer, J. & Knobe, J. 2012 "Contrastive Knowledge Surveyed" in *Nous* Vol. 46 (4). pp. 675-708

Schiffer, S. 2007, "Interest-Relative Invariantism" in *Philosophy and Phenomenological Research* Vol. 75. pp. 188-19

Schwarz, N. *et al,* 1983, "Mood, Misattribution, and Judgments of Well-being" in *Journal of Personality and Social Psychology* Vol. 45. pp. 513-523

Schwarz, N. *et al*, 1991, "Ease of Retrieval as Information" in *Journal of Personality and Social Psychology* Vol. 61. pp. 195-202

Seelu, E. *et al*, 1995, "Counterfactual constraints" in *What Might Have Been: The Social Psychology of Counterfactual Thinking,* Roese, N & Olson, J (*eds*.) (Erlbaum)

Segura, S. *et al,* 2002, "Temporal and causal order effects in thinking about what might have been" in *The Quarterly Journal of Experimental Psychology* Vol. 55A (4). pp. 1295-1305

Sherman, S. *et al*, 1985, "Imagining can Heighten or Lower the Perceived Likelihood of Contracting a Disease" in *Personality and Social Psychology Bulletin* Vol. 111. pp. 118-127

Shoemaker, S. 1994, "Moore's Paradox and Self-knowledge" in *Philosophical Studies* Vol. 77 (2). pp. 211-228

Silins, N. 2005. "Deception and Evidence" in *Philosophical Perspectives* Vol. 19. pp. 375-404

Simon, H. 1955. "A Behavioral Model of Rational Choice" in *Quarterly Journal of*

Economics Vol. 69. pp. 99-118

Simon, H. 1983, *Reason in Human Affairs* (Stanford: Stanford University Press)

Slote, M. 1979, "Assertion and belief" in J. Dancy (*ed*.) *Papers on language and logic*. (Keele: Keele University Library)

Slovic, P. *et al*, 1982, "Facts vs. Fears" in *Judgment under Uncertainty: Heuristics and Biases*, Kahneman. D*, et al* (*eds*.) pp. 463-492. (Cambridge University Press)

Smith, M. 2010, "What else justification could be" in *Nous* Vol. 44 (1). pp. 10-31

Smithies, D. 2012a, "Moore's Paradox and the Accessibility of Justification" pp. 273-300 in *Philosophy and Phenomenological Research* Vol. 85 (2)

Smithies, D. 2012b, "The normative role of knowledge" in *Nous* Vol. 46 (2) pp. 265-288

Sorensen, R. 1988*, Blindspots* (Oxford: Clarendon Press)

Sosa, E. 1999, "How must knowledge be modally related to what is know?" in *Philosophical Topics* Vol. 26 (1&2). pp. 373-384

Srinivasan, A. forthcoming, "Are We Luminous?" in *Philosophy and Phenomenological Research*

Srinivasan, A. manuscript, "What's in a norm?"

Stanley, J. 2005, *Knowledge and Practical Interests* (Oxford University Press: Oxford)

Stanley, J. & Sripada, C. 2012. "Empirical Tests of Interest-Relative Invariantism" in *Episteme* Vol. 9 (1). pp. 3-26

Stanovich, K. 2009, "Distinguishing the reflective, algorithmic, and autonomous minds: Is it time for a tri-process theory?" in *In Two Minds: Dual Processes and Beyond*, Frankish. K, & Evans. J, (*eds*.) pp. 55-88, (Oxford University Press)

Stein, E. 1996, *Without good reason: The rationality debate in philosophy and cognitive science* (Oxford University Press)

Sutton, J. 2005, "Stick To What You Know" in *Nous* Vol. 39 (3). pp. 359-396

Sutton, J. 2007, *Without Justification* (Cambridge, MA: MIT Press)

Turri, J. 2011, "The express knowledge account of assertion" in *Australasian Journal of Philosophy* Vol. 89. pp. 37-49

Unger, P. 1971, "A defense of skepticism" in *Philosophical Review* Vol. 80 (2). pp. 198-219

Unger, P. 1975, *Ignorance* (Oxford: Clarendon Press)

Weatherson, B. 2005, "Can we do without pragmatic encroachment?" in *Philosophical Perspectives* Vol. 19. pp. 417-43

Weatherson, B. 2011, "Defending Interest-Relative Invariantism" in *Logos & Episteme* Vol. 2. pp. 591-609

Weatherson, B. 2012, "Knowledge, Bets, and Interests" in Brown, J & Gerken, M (*eds.*) *Knowledge Ascriptions* pp. 75-103 (Oxford: Oxford University Press)

Weatherson, B. manuscript, "E≠K"

Wedgewood, R. 2008, "Contextualism About Justified Belief" in *The Philosopher's Imprint* Vol. 8 (9). pp. 1-20

Williamson, T. 1996, "Knowing and asserting" in *Philosophical Review* Vol. 105. pp. 489-523

Williamson, T. 2000, *Knowledge and its Limits*, (New York: Oxford University Press)

Williamson, T. 2005, "Contextualism, subject-sensitive invariantism and knowledge of knowledge" in *The Philosophical Quarterly* Vol. 55 (219). pp. 213-235

Williamson, T. 2009, "Reply to Hawthorne and Lasonen-Aarnio" in Greenough, P. & Pritchard, D. (eds.) *Williamson on Knowledge* (Oxford: Oxford University Press)

Williamson, T. 2013, "Response to Cohen, Comesana, Goodman, Nagel, and Weatherson on Gettier cases in Epistemic Logic" in *Inquiry: An Interdisciplinary Journal of Philosophy* Vol. 56 (1). pp. 77-96

Williamson, T. 2014, "Very improbable knowing" in *Erkenntnis* Vol. 79 (5). pp. 971-999

Zagzbeski, L. 1994, "The inescapability of Gettier problems" in *Philosophical Quarterly* Vol 44. pp. 65-73